RUSI

Royal United Services Institute
for Defence and Security Studies

Emerging Insights

Russia, AI and the Future of Disinformation Warfare

Claudia Wallner with Simon Copeland and Antonio Giustozzi

June 2025

EXECUTIVE SUMMARY

As generative AI technologies rapidly evolve, their implications for global information security are becoming more acute. This paper explores how Russian state-affiliated and state-aligned actors are discussing, conceptualising and framing AI within their online communications. Drawing on original analysis of communications from Russian-linked online channels, the paper investigates how actors in the Russian influence ecosystem perceive the role of AI in information warfare and what their narratives reveal about evolving threat trajectories.

The report finds that a diverse range of Russian actors are actively engaged in conversations about AI. These actors are not only discussing the use of AI tools to automate and amplify content, but also exploring the role of AI as a narrative device, boasting of its effectiveness, warning of its dangers and framing it as both a strategic asset and a potential threat.

The analysis reveals a growing focus on AI as both an opportunity and a threat among various Russian actors, from those affiliated with groups like Wagner, to pro-Russian hacktivist collectives and online influencers. AI is often portrayed as a powerful tool for information manipulation, capable of generating persuasive content, amplifying messaging and overwhelming adversaries with sheer volume. At the same time, many actors express significant anxiety about Western dominance over AI development, suggesting that these technologies could be used to subvert Russian public opinion, erode autonomy and destabilise the domestic information environment. Concerns about surveillance, deepfakes (digitally altered videos or images aiming to misrepresent a person as doing or saying something they did not say or do in the original version of the image or video) and algorithmic bias feature prominently in this discourse.

The observed conversations are not confined to abstract speculation. The paper documents how state-affiliated and state-aligned actors are actively debating the implications of AI, sharing practical knowledge, critiquing disinformation practices and recruiting individuals with relevant technical skills. These insights point to an evolving culture of adaptation within Russian influence networks, where AI is increasingly seen as a central component of future-facing information operations.

While the paper does not assess the inner workings of senior intelligence planning, it offers a unique actor-level perspective on how AI is entering the strategic imagination of Russian influence networks. These insights highlight the importance of not only tracking how AI might be operationalised in future disinformation efforts, but also understanding the ways in which it is already shaping how these actors think, communicate and position themselves within digital ecosystems.

INTRODUCTION

The proliferation of generative AI has raised widespread concern over the potential exploitation of AI by malicious actors to disrupt information ecosystems. While early warnings about AI-generated disinformation primarily focused on its theoretical risks, recent evidence suggests that a range of actors are already incorporating these technologies into their influence operations. Understanding how such actors perceive, experiment with and deploy generative AI is essential to anticipating future threats and designing effective countermeasures.

This issue is particularly pressing in the context of Russian information operations. The Russian government has long prioritised information warfare as a central element of statecraft, viewing the information domain as a theatre of conflict on par with conventional or nuclear warfare.[1] Disinformation actors affiliated with the Russian state are thought to have invested heavily in AI technologies to influence European audiences in the run-up to the 2024 European Parliament elections.[2] As generative AI becomes more accessible and more powerful, it lowers the barrier to entry for a wider ecosystem of pro-Russian actors, including state-linked media, hacktivists and online influencers, to experiment with and operationalise these tools in increasingly sophisticated ways.

Despite the severity of this threat, understanding of disinformation and influence operations carried out by Russian state-affiliated actors remains limited. Current discussions focus on outputs, with less attention to how Russian-affiliated actors perceive or discuss the role of AI in influence campaigns. This paper seeks to address this gap by examining how Russian state-affiliated and state-aligned actors and groups – including state-linked social media groups and channels, hacktivist collectives, military-affiliated groups and online influencers – are discussing AI. The paper considers how AI might be a beneficial tool in their arsenal or supplement their existing techniques, tactics and activities.

This paper generates new insight into Russian disinformation and influence activities by scoping and analysing the online communications channels and information ecosystems of these actors and groups. In doing so, it analyses how actors within these spaces discuss AI, including their perceptions, understandings and knowledge of how this technology is currently used, and could be used for propaganda purposes. The paper presents a unique

1. R P Koshkin, 'Artificial Intelligence and Cybernetic Threats to Russia's National Security in Modern Conditions' (author translation), *Культура и безопасность*, 27 February 2020, <https://sec.chgik.ru/en/artificial-intelligence-and-cybernetic-threats-national-security-of-russia-in-modern-conditions/>, accessed 8 April 2025.
2. Recorded Future, '2024 Annual Report,' Cyber Threat Analysis report, 28 January 2025, <https://go.recordedfuture.com/hubfs/reports/cta-2025-0128.pdf>, p.15, accessed 8 April 2025.

perspective on the current and future practices of Russian state-affiliated and state-aligned groups and their attempts to influence and polarise audiences and public opinions, with a particular focus on Europe.

This analysis is situated within a growing body of literature that explores AI-enabled influence operations, while offering a more actor-centric perspective. Rather than treating Russian disinformation as a monolithic, state-directed enterprise, the paper highlights the heterogeneity of actors involved and the diverse, and sometimes contradictory, ways in which they engage with AI.

While the paper focuses on tactical and mid-level actors, it does not attempt to map the strategic or doctrinal thinking of the intelligence managers or decision-makers responsible for designing operations at the highest levels. Such insights are beyond the scope of open-source social media monitoring. Instead, the report contributes to a more granular understanding of how AI is being perceived, discussed and operationalised by those within the broader Russian influence ecosystem, and how these evolving practices reflect emerging threats at the operational and narrative level.

METHODOLOGY

This paper employs a qualitative research design to explore how Russian state-affiliated and state-aligned actors and groups discuss and deploy generative AI in the context of influence operations. The analysis draws primarily on data collected through ExTrac AI's state-of-the-art AI-powered platform.[3] ExTrac AI enables the automatic collection, refinement and analysis of data and multimedia materials from tens of thousands of hard-to-reach but publicly available sources daily, including communications ecosystems associated with disinformation actors.

The research was carried out in three phases, starting with a scoping phase to identify key actors, channels and keywords related to AI and disinformation. Using ExTrac AI's search capabilities, relevant channels were selected based on existing research on Russian influence operations, prior tagging on the platform, and results from initial keyword searches. Keywords included terms in both Russian and English. This stage provided an initial mapping of the online ecosystem of actors engaging with AI and set the foundation for subsequent analytical stages.

In the second stage, content analysis was conducted across a broader dataset, including searches over different time periods and across multiple groups. The terms and themes identified in the scoping phase were used to assess the prevalence and scope of AI-related discourse. Posts were examined to evaluate the degree of interest in AI, the level of technical or

---

3.   'Extrac: Ahead of the Threat', ExtracAI, <https://www.extrac.ai>, accessed 19 June 2025.

operational knowledge demonstrated and the types of narratives emerging around the use of AI for disinformation purposes.

Based on insights from the initial analysis, specific case studies were then selected for further in-depth examination. These case studies focused on particular actors (such as Wagner-affiliated channels and hacktivist groups like NoName057(16)) and explored how AI was discussed and applied within their disinformation efforts.

Together, these three stages provided a structured and iterative approach to understanding the perceptions, use cases and strategic considerations of AI within Russian-aligned disinformation networks.

RUSSIAN STRATEGIC THINKING ON INFORMATION WARFARE AND AI

To assess current and emerging disinformation threats, it is essential to understand how Russian state and military doctrine conceptualises information warfare, and how AI fits within this strategic framework. This section outlines: the foundations of Russian thinking on information confrontation and AI; actors active in this space; and operational practices integrating AI into disinformation operations.

STRATEGIC FOUNDATIONS OF RUSSIAN INFORMATION WARFARE

For Russia, information warfare, or 'information confrontation' as it is known in Russian doctrine,[4] holds strategic importance which is equivalent to conventional and nuclear warfare, even in peacetime.[5] The information domain is treated as integral to national security and geopolitical influence. While Russia rarely acknowledges cyber operations against Western targets, it also avoids directly denying them, using plausible deniability as a strategic asset.[6]

Russian information strategy prioritises intelligence collection, often through espionage or cyber intrusion, alongside information-psychological

---

4. Elina Treyger, Joe Cheravitch and Raphael S Cohen, 'Russian Disinformation Efforts on Social Media', RAND Corporation, 7 June 2022, p. 1, <https://www.rand.org/pubs/research_reports/RR4373z2.html>, accessed 8 April 2025.
5. Blagovest Tashev, Michael Purcell and Brian McLaughlin, 'Russia's Information Warfare: Exploring the Cognitive Dimension,' *MCU Journal* (Vol. 10, No. 2, Fall 2019), pp. 129–47.
6. Koshkin, 'Artificial Intelligence and Cybernetic Threats to Russia's National Security in Modern Conditions'.

influence[7] targeting individuals and mass audiences.[8] This dual focus enables the systematic collection of data while also strengthening efforts to shape political discourse and decision-making domestically and abroad.

Russian disinformation campaigns aim to undermine adversaries by exacerbating internal divisions, eroding trust in democratic institutions, and weakening alliances such as NATO or the EU. This complicates unified responses to Russian actions.[9]

Although many of these campaigns are under-resourced and disorganised, social media platforms allow for low-cost, large-scale experimentation without significant consequence for failed attempts.[10] Trial and error approaches carry little risk, and the volume of content often matters more than precision. Ironically, Western media portrayals of Russian cyber capabilities may amplify their perceived effectiveness, reinforcing Moscow's intended image as a highly sophisticated cyber power, and thereby amplifying the perceived threat even when the operational impact is limited.[11]

The Russian disinformation ecosystem is diverse, encompassing official state entities, state-funded media outlets, proxy organisations, ideologically-driven individuals and commercially-motivated journalists and bloggers.[12] Some operate within formal command structures, while others emerge organically, driven by personal initiative but aligning with perceived Kremlin goals. This decentralised approach provides the Russian state plausible deniability while enabling tailored and flexible messaging that resonates with different audiences without the need for a unified narrative, as would be the case with clearly attributed government communications.[13] A clear example of this decentralised model is evident in Wagner's information operations in Africa, where the group intentionally outsourced its messaging to local activists, journalists and online influencers. These actors are often not directly controlled by Wagner but align with it ideologically or are

> Western media portrayals of Russian cyber capabilities may amplify their perceived effectiveness, reinforcing Moscow's intended image as a highly sophisticated cyber power

7.   Informational and psychological influence refers to the deliberate attempt to shape an individual's or group's thoughts, feelings, and behaviours by manipulating the information they receive.

8.   Андрей Манойло, 'Чему нас научил опыт информационно-психологических операций в ходе СВО' ['What the Experience of Information and Psychological Operations During the SMO has Taught'], Регнум, 19 February 2024, <https://regnum.ru/opinion/3867457>, accessed 8 April 2025.

9.   Scott Jasper, *Russian Cyber Operations: Coding the Boundaries of Conflict* (Washington DC: Georgetown University Press, 2020).

10. Treyger, Cheravitch and Cohen, 'Russian Disinformation Efforts on Social Media', p.24.

11. Bilyana Lilly, *Russian Information Warfare: Assault on Democracies in the Cyber Wild West* (Annapolis, MD: Naval Institute Press, 2022), p.153.

12. US Department of State, Global Engagement Center, 'GEC Special Report: Russia's Pillars of Disinformation and Propaganda', August 2020, <https://2017-2021.state.gov/wp-content/uploads/2020/08/Pillars-of-Russia's-Disinformation-and-Propaganda-Ecosystem_08-04-20.pdf>, accessed 8 April 2025.

13. *Ibid.*

incentivised to support its narratives. These actors allow Wagner to avoid overt branding while achieving narrative resonance, political deniability and operational cost-effectiveness in contested media environments.[14]

At the state level, Russian intelligence agencies, particularly Russia's military intelligence service (GRU), play a central role in orchestrating cyber-enabled influence operations.[15] Two of the GRU's most prominent cyber-focused units, 26165 and 74455, commonly referred to in Western cybersecurity reporting as APT28 (Fancy Bear) and APT29 (Cozy Bear), are implicated in cyber and information operations against foreign targets, including election interference, strategic hacking campaigns and social media influence operations.[16] While these advanced persistent threat actors (APTs) conduct technical operations like intrusions or data theft, they rarely serve as the public face of Russian cyber activity. Instead, hacktivist groups often act as intermediaries, repackaging and disseminating information obtained by APTs to obscure attribution.[17] Factors such as the timing of leaks, shared technical infrastructure, and instances where hacktivist groups claim responsibility for an operation before it has been publicly attributed to an APT often suggest coordination.[18] This interplay allows the Russian state to expand influence while diffusing risk and complicating attribution.

Additionally, proxy media outlets such as Strategic Culture Foundation, Global Research, New Eastern Outlook, News Front, SouthFront, Katehon and Geopolitica.ru function as multipliers of Russia's narratives abroad.[19] They extend Russia's messaging reach abroad by reproducing and legitimising disinformation narratives without requiring direct Kremlin oversight.

AI AS AN ENABLER OF DISINFORMATION

Russia has identified AI as a strategic domain.[20] Russia's 2019 national strategy for the development of AI committed the country to global leadership in AI

---

14.   Antonio Giustozzi, 'What Next for Wagner's Information Operations?', *RUSI Commentary*, 23 September 2024.

15.   Lilly, *Russian Information Warfare*, p. 27.

16.   Treyger, Cheravitch and Cohen, 'Russian Disinformation Efforts on Social Media'; Bundesamt für Sicherheit in der Informationstechnik, 'Aktive APT-Gruppen in Deutschland' ['Active APT groups in Germany'], <https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Cyber-Sicherheitslage/Analysen-und-Prognosen/Threat-Intelligence/Aktive_APT-Gruppen/aktive-apt-gruppen_node.html>, accessed 8 April 2025.

17.   Treyger, Cheravitch and Cohen, 'Russian Disinformation Efforts on Social Media'.

18.   Insikt Group, 'Themes and Failures of Russia's War Against Ukraine', 9 February 2023, <https://www.recordedfuture.com/research/themes-failures-russias-war-against-ukraine>, accessed 8 April 2025.

19.   US Department of State, 'Pillars of Russia's Disinformation and Propaganda Ecosystem'.

20.   Radina Gigova, 'Who Vladimir Putin Thinks Will Rule the World', CNN, 2 September 2017.

by 2030.[21] Russian officials frame AI as both an opportunity and a threat.[22] Officials express concern that Western AI tools could manipulate Russian public opinion, introduce ideological bias, or even catalyse regime change.[23] These concerns reinforce Russia's belief in AI's capacity to enable cognitive manipulation and encourage internal destabilisation.[24]

In response, the Kremlin has prioritised a sovereign AI ecosystem, coordinated by the National Centre for AI Development.[25] State-owned enterprises, such as Sberbank and Rostec, dominate the field.[26] Sberbank's GigaChat platform, which is presented as Russia's answer to OpenAI's ChatGPT, is a flagship initiative, serving as both a technological benchmark and an instrument for state AI policy – with Sberbank being tasked with implementing Russia's AI roadmap and the national strategy for the development of AI. Rostec focuses on defence applications, while Yandex (creator of YandexGPT) – despite its leading AI capabilities – plays a secondary role due to its complicated relationship with the Kremlin.[27]

Generative AI is already being integrated into Russian disinformation operations. Automated tools generate fake articles, social media posts, images and deepfakes.[28] Operations like the 'DoppelGänger' campaign, in which AI-generated articles mimicked legitimate Western news outlets, illustrate how these tactics aim to erode trust and sow confusion at scale.[29] AI-powered bots and automated social media accounts help amplify disinformation, saturate public discourse and simulate grassroots sentiment

21. Hybrid Warfare Analytical Group, 'Artificial Intelligence in the Kremlin's Information Warfare', Ukrainian Crisis Media Center, 20 February 2025, <https://uacrisis.org/en/artificial-intelligence-in-the-kremlin-s-information-warfare>, accessed 8 April 2025.

22. Samuel Bendett, 'The Role of AI in Russia's Confrontation with the West', Center for a New American Security, Transatlantic Security Program, 3 May 2024, p.7, <https://www.cnas.org/publications/reports/the-role-of-ai-in-russias-confrontation-with-the-west>, accessed 8 April 2025.

23. Hybrid Warfare Analytical Group, 'Artificial Intelligence in the Kremlin's Information Warfare'.

24. Bendett, 'The Role of AI in Russia's Confrontation with the West', p.19.

25. Ukraine Crisis Media Center, 'Artificial Intelligence in the Kremlin's Information Warfare'.

26. Stephanie Petrella, Chris Miller and Benjamin Cooper, 'Russia's Artificial Intelligence Strategy: The Role of State-Owned Firms', *Orbis* (Vol. 65, No. 1, 2021), pp. 75–100.

27. *Ibid*.

28. Karen Allen and Christopher Nehring, *AI-Generated Disinformation in Europe and Africa: Use Cases, Solutions and Transnational Learning* (Johannesburg: Konrad-Adenauer-Stiftung Media Programme Sub-Saharan Africa, 2025), p.89.

29. USCYBERCOM Public Affairs, 'Russian Disinformation Campaign "DoppelGänger" Unmasked: A Web of Deception', 3 September 2024, <https://www.cybercom.mil/Media/News/Article/3895345/russian-disinformation-campaign-doppelgnger-unmasked-a-web-of-deception/>, accessed 8 April 2025.

– a tactic known as 'astroturfing'.[30] In some cases, fake conversations between bots are staged to simulate debate and mislead third-party observers.[31]

Russian actors are also exploring large language model (LLM) grooming: injecting of propaganda or biased material into training data to influence the outputs of LLMs. Designed to skew the outputs of LLMs, this tactic represents a shift from targeting audiences directly to subtly shaping the tools these audiences use.[32]

Though the use of generative AI in disinformation is still evolving, it is increasingly seen by Russian actors as a force multiplier, automating content production, overwhelming adversaries and further blurring the boundary between truth and fabrication.

MAPPING AI-RELATED DISCOURSE IN RUSSIAN INFLUENCE NETWORKS

This section draws on primary data gathered through the ExTrac AI platform to examine how Russian state-affiliated and state-aligned actors discuss, conceptualise and operationalise generative AI technologies.

KNOWLEDGE OF AND FAMILIARITY WITH AI TOOLS

Analysis of the discussions under review reveals a strong and growing interest among Russian actors in the capabilities of AI, particularly as a tool in disinformation campaigns and broader strategic communication. While much of the discourse references the military applications of AI, especially the integration of AI with drone operations, there are frequent and direct acknowledgements of its value for information operations and manipulation.

Numerous Telegram channels express pride in sophisticated Russian information operations, explicitly citing the use of AI-driven tools. Some channels even frame AI as central to training a new generation of cyber operatives, which reflects their intent to further establish and institutionalise AI within the disinformation apparatus.[33]

Western media portrayals of Russian cyber capabilities may amplify their perceived effectiveness, reinforcing Moscow's intended image as a highly sophisticated cyber power

30. O E Voronova, *Sovremennye informatsionnye voiny: tipologiya i tekhnologii: monografiya* [Modern Information Wars: Typology and Technologies] (Ryazan: Ryazan State University, 2018), p. 43.
31. O E Voronova, *Sovremennye informatsionnye voiny: tipologiya i tekhnologii: monografiya*, pp. 87, 153.
32. VIGINUM, 'PORTAL KOMBAT: A Structured and Coordinated Pro-Russian Propaganda Network', Secrétariat général de la défense et de la sécurité nationale, technical report, February 2024, <https://www.sgdsn.gouv.fr/files/files/20240212_NP_SGDSN_VIGINUM_PORTAL-KOMBAT-NETWORK_ENG_VF.pdf>, accessed 8 April 2025.
33. Telegram, collected via ExTrac AI, 24 February 2022; Telegram, 7 January 2025.

Alongside bravado and dismissiveness, the observed discussions also reveal a certain level of underlying distrust and paranoia about AI technologies

At the same time, these actors also mock and ridicule Western media coverage of Russian AI-driven interference. For example, in response to Western media reports accusing Russian actors of running German-language influence campaigns, some posts sarcastically questioned who in Russia could possibly speak German well enough to do so, humorously implying direct involvement of Vladimir Putin himself, given his German language skills.[34] Actors dismiss reports from German or other European news outlets about Russian cyber and AI-driven influence operations as exaggerated or fabricated, portraying these accusations as attempts by Western governments to deflect attention from domestic political failures.[35] Such derisive dismissals aim to undermine the credibility of Western reports while reinforcing an internal narrative that Russia's influence operations effectively unsettle Western governments.

Yet, alongside bravado and dismissiveness, the observed discussions also reveal a certain level of underlying distrust and paranoia about AI technologies. A range of concerns were expressed, from existential anxieties about the future capabilities of AI to practical fears about the immediate implications of AI technologies.[36] For instance, posts expressed anxieties regarding the ability of AI to rapidly and convincingly reproduce human personalities, seeing this as a dystopian threat capable of gradually replacing humans altogether.[37] Another fear expressed involves the potential creation of artificial superintelligence (ASI),[38] with actors describing scenarios in which routine human activities are rapidly replaced by AI agents and predicting that AI systems will achieve human-level intelligence, eventually culminating in an entirely AI-controlled society.[39]

Other concerns are more immediate and pragmatic. For example, the inclusion of the former US NSA director, Paul Nakasone, on OpenAI's board was widely interpreted as proof of the ongoing militarisation of AI. Actors suggested that all interactions with OpenAI systems could now (potentially) be subject to surveillance and exploitation by US military intelligence.[40] These narratives demonstrate both familiarity with ongoing AI developments in the West and heightened anxiety about losing control over these technologies.

In this context, Ukraine is often framed by channels as a testing ground for AI-enabled psychological influence operations, mass surveillance and propaganda by Western intelligence agencies.[41] For example, channels point to supposedly AI-generated intercepted communications between

---

34. Telegram, collected via ExTrac AI, 26 January 2024.
35. Telegram, collected via ExTrac AI, 26 January 2024; Telegram, collected via ExTrac AI, 19 March 2024; Telegram, collected via ExTrac AI, 4 June 2024.
36. Telegram, collected via ExTrac AI, 15 September 2024.
37. Telegram, collected via ExTrac AI, 8 January 2025.
38. ASI refers to a hypothetical AI that surpasses human intelligence in all aspects, including creativity, problem-solving, and emotional intelligence.
39. Telegram, collected via ExTrac AI, 24 July 2024.
40. Telegram, collected via ExTrac AI, 15 June 2024.
41. Telegram, collected via ExTrac AI, 18 January 2025

Russian soldiers and their relatives, deepfake videos falsely attributed to Russian officials, and manipulative social media campaigns, arguing that these represent coordinated psychological operations designed to weaken Russian resolve and morale.[42]

Moreover, specific examples like the DoppelGänger campaign are framed ambiguously, with some channels insisting such campaigns represent false flags orchestrated by Western intelligence in collaboration with Russian opposition figures, thus reinforcing narratives of victimhood and external aggression.[43]

Such concerns have led Russian-aligned actors to call explicitly for improved information literacy among the Russian public. Frequent warnings are posted urging vigilance against perceived Western deepfake operations and manipulative content. Channels regularly publish reminders for followers to verify information rigorously, while remaining sceptical of unofficial sources and relying primarily on state-sanctioned media outlets.[44]

FRAMING AND FUNCTION OF AI TOOLS

Within the online communities under review, AI tools are framed as essential components in an intensifying information conflict. Telegram channels present AI as not only a technological innovation but a civic duty, encouraging skilled individuals to contribute to national efforts. Recruitment calls often target AI-literate users to contribute their skills to the groups' goals, in line with Russia's perceived interests.

For example, one Telegram channel actively recruits contributors with experience in generative AI tools, stating: 'Are you a neural network artist? [Redacted: name of channel] needs you! Our team is developing and the number of tasks is growing every month. We are looking for artists already working with Stable Diffusion (AUTOMATIC1111+ControlNet) and Midjourney ... who understand the principles of industrial engineering and can control generation rather than rely on luck.'[45] Similarly, another recruitment message from the same channel details its broad range of desirable recruits, including 'neuroevangelists and text generation specialists', emphasising that the project's activities extend beyond graphics into sophisticated textual manipulation and generation.[46] The channel describes its project as fundamentally reliant on human capital – including talented analysts, informants, and digital intelligence enthusiasts – highlighting collaborative 'brainstorming' and diverse expertise spanning geospatial intelligence and

---

42.   Telegram, collected via ExTrac AI, 26 December 2024.

43.   Telegram, collected via ExTrac AI, 9 March 2024.

44.   Telegram, collected via ExTrac AI, 7 August 2024; Telegram, collected via ExTrac AI, 10 August 2024; Telegram, collected via ExTrac AI, 16 January 2025; Telegram, collected via ExTrac AI, 31 January 2025.

45.   Telegram, collected via ExTrac AI, 22 August 2024.

46.   Telegram, collected via ExTrac AI, 17 March 2024.

open-source intelligence.[47] The channel actively appeals to individuals who can navigate complex international information environments and generate persuasive narratives which are supportive of Russian military and geopolitical objectives.[48]

Another central topic in the observed discourse is the centrality of the information war as a part of contemporary conflict and national security. For instance, promotional materials for the 'Army-2024' Forum's roundtable[49] emphasised the importance of understanding how advanced digital technologies, including AI, can shape societal perceptions in the context of Russia's 'Special Military Operation' in Ukraine.[50]

Other Telegram channels frame AI tools as instrumental resources for safeguarding Russia's digital space against external threats. For example, one channel under review describes itself as a community of 'highly qualified specialists in cybersecurity, information technology, and social research', dedicated explicitly to 'neutralising threats, disinformation, and propaganda'.[51] These discussions portray AI as a tool of not only offence but also defence, used to identify, counter and neutralise foreign influence operations allegedly targeting Russian society.

PERCEIVED LIMITATIONS AND BARRIERS

Despite significant enthusiasm surrounding the strategic integration of AI into information warfare, the monitored online communities also expressed substantial criticism and frustration regarding the limitations of Russia's domestic AI platforms, primarily Sberbank's GigaChat and YandexGPT. These criticisms reflect broader anxieties about technological autonomy, ideological biases and operational constraints, as well as suspicions regarding the political loyalties of major Russian tech firms (particularly Yandex). A recurring critique is that these models reflect liberal or 'unpatriotic' biases.[52] For instance, Sber GigaChat has been accused by users of showing favourable attitudes toward figures such as Lenin and Trotsky; one post complained that the platform 'cannot condemn their betrayal of Russia and the Red Terror'.[53]

Even more concerning for these actors is the platforms' handling of politically sensitive topics, especially those involving Russian territorial claims. YandexGPT and Sber GigaChat have both faced accusations of failing

---

47. *Ibid.*

48. *Ibid.*

49. The roundtable was titled 'Информационное противоборство в условиях СВО: борьба за коллективное сознание общества с использованием передовых технологий и цифровых СМИ' ['Information Confrontation in the Context of the Special Military Operation: The Battle for the Collective Consciousness of Society Using Advanced Technologies and Digital Media'], (Extrac AI translation).

50. Telegram, collected via ExTrac AI, 14 August 2024.

51. Telegram, collected via ExTrac AI, 19 April 2024.

52. Telegram, collected via ExTrac AI, 18 January 2025.

53. *Ibid.*

to confirm that Crimea and the recently annexed 'new regions' are part of Russia.[54] Posts shared anecdotes in which the AI would either avoid such questions or suggest changing the topic, which was interpreted by users as evasive and subversive. One user sarcastically noted that 'Sber GigaChat is starting to freeze even when asked questions like "which regions by the sea would Russians like to move to" (because Crimea is on the list)'.[55] While later Telegram discussions noted an improvement, highlighting that 'Gigachat already cheerfully answers that Crimea and Sevastopol have become part of Russia',[56] Yandex continues to be accused of intentionally misrepresenting sensitive facts about Ukraine as well as misrepresenting historical events.

This reluctance to reproduce state-endorsed positions and narratives has led to accusations that platforms like YandexGPT are either deliberately undermining Russia or are bound to foreign interests – a matter that users think should be escalated to the level of Russia's Security Council for intervention.[57] As one Telegram post stated: 'Yandex GPT and Alisa still cannot be forced to confirm the constitutional norm about Russian Crimea … Russian developers are not so Russian, and the enemy who is waging the main – mental – war with us is not outside, but inside our country'.[58]

Other domestic platforms have faced similar allegations. Megafon's chatbot, for instance, was criticised for classifying Crimea and the Donetsk and Luhansk People's Republics as Ukrainian territories, allegedly because the underlying AI architecture relies on foreign-developed technologies such as Midjourney and GPT.[59]

Beyond ideological concerns, practical limitations possessed by Russian tools, when compared to Western tools such as ChatGPT, were another major source of frustration discussed in the observed channels.[60] Users complained that Russian services were less responsive, provided inferior responses or declined to address neutral and factual questions that similar Western tools would answer. One post elaborated, saying, 'let's take YandexGPT … this prototype of AI is a terrible coward and does not answer quite ordinary questions … On the one hand, this greatly undermines trust in Yandex and its products. On the other hand, it gives grounds not only to recognize Yandex's services as incomplete, but even [to recognise] the current managers … as foreign agents.'[61]

These perceived shortcomings undermine Russia's stated goals of achieving digital and technological sovereignty. Ironically, pro-Russian

54.   Telegram, collected via ExTrac AI, 19 May 2024.
55.   Telegram, collected via ExTrac AI, 8 November 2024.
56.   Telegram, collected via ExTrac AI, 8 July 2024.
57.   Telegram, collected via ExTrac AI, 19 May 2024.
58.   Telegram, collected via ExTrac AI, 8 November 2024.
59.   Telegram, collected via ExTrac AI, 8 July 2024.
60.   Telegram, collected via ExTrac AI, 7 December 2024.
61.   Telegram, collected via ExTrac AI, 5 July 2024.

actors' dissatisfaction with domestic AI has led to a continued preference for Western tools, which undermines the Kremlin's narrative of building a successful autonomous and ideologically-aligned AI infrastructure.

CASE STUDIES

To better understand how these discussions around AI translate into practical influence operations, this section presents an analysis of two key actor groups: Wagner-affiliated channels and Russian state-aligned hacktivist collectives.

WAGNER-AFFILIATED CHANNELS

Wagner, the Russian state-funded private military company known for its roles in Ukraine, including its early involvement in the Donbas war, its recruitment of prison inmates, and its central role in the Battle of Bakhmut, has also supported authoritarian regimes in Syria, Libya, the Central African Republic, and Mali.[62] In Africa, its operations have often involved exchanging security services for access to resource concessions, particularly in the mining sector. The group has been repeatedly accused of serious human rights abuses, including torture, extrajudicial killings, and war crimes.[63]

In parallel with its military activity, Wagner is also highly active in the information warfare space. Channels affiliated with Wagner on Telegram and other semi-public platforms reveal significant engagement with generative AI technologies. They position these technologies as tools to undermine trust in Western institutions, sow discord among populations in the West and frame any Russian cyber activities as defensive responses to perceived Western aggression.

Many of the reviewed discussions in Wagner-linked channels contained critiques of poorly organised and ineffective Russian attempts at information warfare that relied on impulsive actions rather than systematic, intelligence-driven methods.[64] This included, for example, criticism of low quality, AI-generated images and videos intended to blur reality. Critics argued that poorly executed efforts damage credibility and limit the effectiveness of genuine operations; as one post put it: 'Information confrontation is

---

62. Antonio Giustozzi, Joana de Deus Pereira and David Lewis, 'Did Wagner Succeed in the Eyes of its African and Middle Eastern Clients?' *RUSI Whitehall Report*, 9 January 2025.

63. It is important to acknowledge the group's turbulent trajectory following its failed insurrection against the Russian government in June 2023. The fallout from that episode led to a reassertion of state control over Wagner's operations and a partial absorption of its assets into formal Russian structures, raising questions about the future coherence and centrality of Wagner-linked influence networks.

64. Telegram, collected via ExTrac AI, 21 June 2024.

necessary, but it must be done wisely – not in such a way as to report to the top for the sake of a tick'.[65] This commentary reflects Wagner's framing of itself as a disciplined, professional actor that values careful, intelligence-based planning – as opposed to careless disinformation and impulsive or amateur operations (which, they argue, risk undermining broader strategic efforts). By portraying other actors as ineffective, Wagner reinforces its own claim to be a capable and credible partner in Russia's information warfare architecture.

To counter enemy influence in the information space and build up resilience among Wagner's operatives and audiences, Wagner-linked channels promote education in psychological operations and information warfare. They advertise courses and certifications offered by Wagner-affiliated figures designed to help their supporters identify enemy psychological tactics, detect fake content and improve strategic communications literacy. One teaching aid distributed in early 2025 stated: 'I tried to collect the most common methods the enemy uses daily … this teaching aid briefly outlines the essence of the information war and the main methods for recognising fake materials. After studying this manual, you will learn to recognise and counter manipulation.'[66]

Wagner-linked channels also actively encourage their members to familiarise themselves with ongoing AI developments, highlighting, for example, advancements toward artificial general intelligence (AGI)[67] and ASI, and discussing the potential strategic implications of these emerging technologies. They share practical information about new AI tools such as Google Labs' image generation tool Whisk, explaining its potential operational applications and highlighting how AI-generated content can be strategically used in their activities.[68] At the operational level, Wagner channels have discussed incorporating AI bots for moderation, as well as exploring the creation of bespoke AI chatbots modelled on popular or authoritative Wagner-affiliated figures.[69]

At the same time, Wagner-linked channels also express warnings about potential adversarial use of AI, particularly highlighting threats posed by AI-generated deepfakes designed to undermine Wagner operations internationally. For instance, Wagner-linked channels alerted followers to attempts to discredit Wagner's presence in the Sahel using AI-generated images.[70] Wagner-linked channels also drew attention to deepfake operations aimed at undermining regional authorities in Kursk,[71] and to AI-generated audio used by Ukrainian forces to psychologically target Russian

---

65.   Telegram, collected via ExTrac AI, 27 May 2024.
66.   Telegram, collected via ExTrac AI, 5 February 2025.
67.   AGI refers to a hypothetical type of AI that can perform any intellectual task that a human can.
68.   Telegram, collected via ExTrac AI, 25 December 2024.
69.   Telegram, collected via ExTrac AI, 13 June 2024.
70.   X, collected via ExTrac AI, 12 January 2024.
71.   Telegram, collected via ExTrac AI, 7 August 2024.

soldiers and their families. As one Wagner-affiliated channel warned: 'A deepfake was posted by Ukrainian TsIPsO [ЦІПСО – referring to Ukraine's Centre for Information and Psychological Operations] ... they trained a neural network quickly with just one short video. The pipeline of deepfakes will soon escalate; remain vigilant.'[72]

Another case discussed in the observed channels involved claims that the French Cyber Defence Command trained Ukrainian and Polish cyber units specifically to target Wagner operations in Mali and elsewhere. These channels reported that the operations were part of a wider Western effort to weaponise AI against Russia and Russian interests.[73] These warnings are simultaneously contributing to a heightened sense of alertness, and amplifying the narrative of being under constant threat by hostile powers – serving to justify campaigns against the West.

Further complicating Wagner's relationship with AI are controversies where AI-generated content has reportedly been misused to manipulate or misrepresent Russian domestic perspectives. In one widely discussed case, a controversial statement, supposedly from State Duma Deputy Alexander Borodai, was dismissed as an AI-generated fake after it described frontline volunteers as 'spare people'.[74] Wagner-linked channels leaned into the controversy by sarcastically noting that the 'neural network Borodai' had simply revealed the unvarnished truth of elite opinion.[75] This ambiguity, blending satire with critique, illustrates Wagner's complex relationship with the Russian state as well as its engagement with AI as both a powerful tool and a potential liability in the struggle over narrative control.

In addition to its operational focus, Wagner's discourse around AI often reinforces a metanarrative of elite competence. Channels often contrast Wagner's disciplined, intelligence-led approach to information warfare with what they describe as amateurish or ideologically confused efforts by other Russian actors. For example, commenting on the Russian information campaign 'Dozor', one channel member commented, 'this initiative will do more harm than good ... the activity is poorly organized, there is no systematic approach ... All these games with telegram bots are not the level of the challenges we face. Until this work is organized and set up properly, it is unlikely that anything will work out'.[76] In this framing, Wagner positions itself as a pioneer of modern Russian information operations: technologically literate, strategically minded, and better equipped to utilise

---

72.   Telegram, collected via ExTrac AI, 26 December 2024.

73.   Telegram, collected via ExTrac AI, 1 January 2025.

74.   Nikita Sologub, '"Spare people". Russian MP and exDPR Leader Alexander Borodai Caught on Tape Describing Russian Volunteers as Expendable in Ukraine War,' Mediazona, 4 November 2024, <https://en.zona.media/article/2024/11/04/borodai-trl>, accessed 11 June 2025.

75.   Telegram, collected via ExTrac AI, 2 November 2024.

76.   Telegram, collected via ExTrac AI, 21 June 2024.

the full potential of AI than other actors who are constrained by bureaucracy or political leanings.

HACKTIVIST GROUPS

Pro-Russian hacktivist collectives – many of which operate either under the direction of, or in strategic alignment with, state intelligence entities such as the GRU, the external intelligence agency (SVR), and the state security agency (FSB) – form another important component of Russia's broader influence and cyber-operations ecosystem. Prominent groups include Zarya, Cyber Army of Russia (also known by other aliases such as the People's Cyber Army or Cyber Army of Russia Reborn[77]), Solntsepek, Beregini and RaHDit, and NoName057(16), and appear to have ties to APT units such as APT44 (GRU Unit 74455).[78]

These groups often target anti-disinformation organisations, independent media and strategic infrastructure, particularly in Ukraine and other parts of Europe, framing these actions as retaliatory strikes against Western 'information aggression'.[79] For example, in May 2024, the Cyber Army of Russia attacked the website of Ukraine's Centre for Combating Disinformation, justifying its actions as retaliation against perceived Ukrainian misinformation about incidents in Belgorod: 'We must make every effort to teach them a lesson in poisoning the information space with cynical lies about yesterday's shelling of a residential building in Belgorod.'[80]

Beyond conducting attacks, these groups also exploit the symbolic value of media attention. For example, distributed denial-of-service (DDoS) campaigns which receive significant Western media coverage, such as the temporary disruption of the official website of Slovenia's president, are showcased as major victories, regardless of their technical or strategic significance.[81] Screenshots of affected websites are circulated as digital trophies that are intended to boost morale among their supporters and signal success, both internally and to the groups' adversaries.[82]

---

77. Cyble, 'Threat Actor Profile: People's Cyber Army of Russia,' 20 March 2025, <https://cyble.com/threat-actor-profiles/peoples-cyber-army-of-russia/>, accessed 18 June 2025.
78. Microsoft Threat Intelligence, 'Russia-Linked Operators Engaged in Expansive Efforts to Influence US Voters,' Microsoft, 27 September 2024, <https://www.microsoft.com/en-us/security/security-insider/intelligence-reports/russia-linked-operators-engaged-in-expansive-efforts-to-influence-us-voters>, accessed 8 April 2025; Mandiant Intelligence, 'Hacktivists Collaborate with GRU-sponsored APT28,' Google Cloud blog, 23 September 2022, <https://cloud.google.com/blog/topics/threat-intelligence/gru-rise-telegram-minions>, accessed 18 June 2025.
79. Telegram, collected via ExTrac AI, 2 April 2024.
80. Telegram, collected via ExTrac AI, 13 May 2024.
81. Telegram, collected via ExTrac AI, 2 April 2024.
82. *Ibid.*

Another important part of these groups' modus operandi is the sharing of knowledge, resources and technical expertise. For instance, they highlight and promote platforms, such as HackerGPT, which are explicitly tailored to aid Russian-aligned hackers and which provide databases of techniques, tools and strategies designed to facilitate effective cyber operations.[83] Other resources disseminated in hacktivist channels include comprehensive repositories on hacking, tutorials on generative AI, and lists of AI tools for different purposes.[84] The aim is to cultivate a technically literate, decentralised cyber force, capable of sustaining long-term operations with minimal central oversight.

The hacktivist group NoName057(16) exemplifies the convergence of AI and cyber operations.[85] Since emerging in early 2022, NoName057(16) has openly discussed AI as a force multiplier for DDoS attacks, misinformation campaigns, and reputational sabotage. The group positions generative AI predominantly as an operational enhancer rather than a standalone capability.

The group demonstrates awareness of international research and reports analysing its influence operations.[86] The channel actively references external reports, such as Google's researchers identifying AI as a principal source of online misinformation,[87] treating them as external validation of their operational effectiveness.[88] This engagement with adversarial reporting and scrutiny blurs the line between propaganda and performance, creating a feedback loop where Western scrutiny is co-opted to reinforce the group's narrative of impact and relevance.

When alleged members of the group were arrested in Spain in mid-2024, NoName057(16) framed the incident as a symptom of a broader European 'witch hunt', driven by what they described as 'Russophobic authorities'.[89] This narrative serves to reinforce group cohesion and present their activities as legitimate resistance against unjustified Western persecution.

---

83. Telegram, collected via ExTrac AI, 2 March 2024.

84. Telegram, collected via ExTrac AI, 27 May 2024

85. Arturo Di Corinto, 'The Role of Disinformation, Propaganda and Active Measures in Cyber Warfare: Noname(057)16 Travels to Italy', presented to *ITASEC 2024: The Italian Conference on CyberSecurity*, Salerno, 8–12 April 2024, published in *CEUR Workshop Proceedings*, Vol. 3731, <https://ceur-ws.org/Vol-3731/paper26.pdf>, accessed 8 April 2025.

86. Telegram, collected via ExTrac AI, 30 August 2024; Telegram, collected via ExTrac AI, 2 May 2024.

87. Maggie Harrison Dupré, 'Even Google's Own Researchers Admit AI Is Top Source of Misinformation Online', Yahoo News, 29 May 2024, <https://nz.news.yahoo.com/even-google-own-researchers-admit-184720725.html>, accessed 11 June 2025.

88. Telegram, collected via ExTrac AI, 30 May 2024.

89. Telegram, collected via ExTrac AI, 22 July 2024.

Finally, NoName057(16) actively cultivates its public identity through interviews,[90] digests, and multi-language media products, positioning itself as a committed actor in Russia's information war. Its Telegram channels call for ongoing cyber operations using AI,[91] provide tools and guides to users,[92] and foster a sense of community built around technical skill and ideological conviction.

CONCLUSION

As this report has shown, generative AI is no longer a concern of the future, but an active component of ongoing Russian-aligned influence operations. A range of actors are already engaging with AI not only as a tool to amplify their content, but also as a conceptual and strategic asset. These actors are debating the potential of AI, educating their followers and supporters in its use, and integrating it into broader narratives of geopolitical struggle and digital sovereignty.

Russian discourse around AI reveals both strategic opportunism and deep insecurity. AI is praised for its ability to scale and personalise disinformation, automate content generation and reduce attribution risks. At the same time, there is widespread concern about the militarisation of Western AI, the ideological bias of foreign-developed tools, and the perceived failure of Russian platforms to align with the state's messaging priorities. This reflects a broader tension in Russia's approach: while seeking to dominate the information space, it remains dependent on digital tools and infrastructures it does not control.

The analysis highlights that generative AI is not simply being used to upgrade existing disinformation techniques. Instead, it is shaping how influence operations are developed, legitimated and operationalised. Wagner's use of AI-enhanced educational materials, its criticism of low-quality content and its call for strategic discipline all suggest a professionalisation of the information war. Meanwhile, hacktivist groups like NoName057(16) emphasise AI as a force multiplier for grassroots cyber operations, embedding it into decentralised campaigns aimed at overwhelming and destabilising Western digital infrastructure.

These trends have several policy and security implications. First, the fusion of AI and influence operations reinforces the need for AI governance frameworks that explicitly address malign use cases, not only in terms of content generation, but also in relation to model training, access and deployment. Second, the decentralised and multilingual nature of Russian-aligned influence networks highlights the importance of monitoring actor discourse, including strategic framing and planning discussions, across various platforms, rather than just focusing on outputs, such as

---

90.   Telegram, collected via ExTrac AI, 13 November 2024.
91.   Telegram, collected via ExTrac AI, 10 January 2025.
92.   Telegram, collected via ExTrac AI, 10 January 2025.

disinformation posts or fake media content. Without this monitoring, the ability to anticipate emerging tactics and narratives is significantly reduced.

There is also a need to support civil society and media ecosystems that are directly targeted by these operations. Aside from protecting them from AI-enabled attacks, investment in digital literacy and resilience programming against synthetic content and manipulated narratives is also necessary. Given the pace of AI development, coordination between governments, platforms, researchers and journalists must also happen at a larger scale. This should include sharing insights on observed tactics and uses of AI tools, as well as behaviours of threat actors.

Ultimately, this report demonstrates that AI is reshaping, but not replacing, the mechanics and logic of Russian disinformation. AI acts as an amplifier, enabling greater reach, faster response and more dynamic narrative adaptation. But it also introduces new vulnerabilities, contradictions and frictions within Russian influence networks themselves. Understanding and engaging with these internal dynamics will be important to inform future policy design and the development of effective countermeasures.

ABOUT THE AUTHORS

Claudia Wallner is a Research Fellow in RUSI's Terrorism and Conflict research group, specialising in the prevention and countering of violent extremism (P/CVE), with an emphasis on far-right extremism and terrorism. She co-leads RUSI's Far-Right Extremism and Terrorism research programme and is part of a team delivering EU-funded P/CVE trainings in different regions.

Claudia's work includes research on the role of gender dynamics in far-right radicalisation and recruitment, as well as the issue of far-right extremism within the security forces. Her work also addresses online dynamics of violent extremism, including the intersection of extremism with online gaming environments, and evaluates the effectiveness of P/CVE interventions in different geographic and thematic areas. Claudia has a Masters with Distinction in Conflict, Security and Development from King's College London and a Masters in Countering Organised Crime and Terrorism from University College London.

Simon Copeland is a researcher specialising in violent extremist narratives and networks, strategic communications and counterterrorism. He was previously a Research Fellow in RUSI's Terrorism and Conflict research group. Prior to RUSI, he worked as a researcher at Swansea University examining online terrorist propaganda, in particular how jihadist and right-wing groups attempt to inspire their supporters. He has also worked as a Research Associate at the Centre for Research and Evidence on Security Threats, where he produced a series of reports on issues related to violent extremism and the UK counter-terrorism strategy (CONTEST). In other roles, he has also analysed terrorist and extremist groups in Sub-Saharan Africa, including in Kenya, Somalia, and Nigeria.

Simon completed his PhD in Politics and International Relations from Lancaster University in 2019, focusing on the role of kin and peer networks in shaping violent extremists' worldviews. He also holds an MA in International Security Studies.

Antonio Giustozzi took his PhD at the London School of Economics and Political Science (LSE) and is currently Senior Research Fellow at RUSI. He is the author of several articles and papers on Afghanistan, as well as of six books: *War, Politics and Society in Afghanistan, 1978-1992* (Georgetown University Press); *Koran, Kalashnikov and Laptop: The Neo-Taliban Insurgency in Afghanistan* (Columbia University Press); *Empires of Mud: War and Warlords in Afghanistan* (Columbia University Press); *Policing Afghanistan: The Politics of the Lame Leviathan* (with M Ishaqzada, Columbia University Press), *The Islamic State in Khorasan* (Hurst) and *The Taliban at War* (Hurst). Dr Giustozzi also edited a volume on the Taliban, *Decoding the New Taliban: Insights from the Afghan Field* (Columbia University Press), featuring contributions by specialists from different backgrounds.

Beyond Afghanistan, Dr Giustozzi published articles on the conflict in Syria and jihadist groups in Central Asia, authored a volume on the role of coercion and violence in state-building, *The Art of Coercion: The Primitive Accumulation and Management of Coercive Power* (Columbia University Press) and edited another volume on disarmament, demobilisation and reintegration processes, *Post-Conflict Demobilisation, Disarmament and Reintegration: Bringing State-Building Back In* (Ashgate).

ACKNOWLEDGEMENTS

PROJECT PARTNER — EXTRAC AI

ExTrac finds, curates, and ingests the highest relevance and hardest-to-reach adversary data sources at scale across the surface, deep, and dark web in a compliant and secure manner. It captures the communications and kinetic behaviours of hostile state and non-state actors and mines them for real-time insights on conflict, greyzone, and influence activities using AI that's both secure and compliant by design.

194 years of independent thinking on defence and security

The Royal United Services Institute (RUSI) is the world's oldest and the UK's leading defence and security think tank. Its mission is to inform, influence and enhance public debate on a safer and more stable world. RUSI is a research-led institute, producing independent, practical and innovative analysis to address today's complex challenges.

Since its foundation in 1831, RUSI has relied on its members to support its activities. Together with revenue from research, publications and conferences, RUSI has sustained its political independence for 194 years.