Centre for International
Governance Innovation

Policy Brief No. 227 — February 2026

From Trolls to Generative AI: Russia's Disinformation Evolution

Daniela Iampolsca

Key Points

- Kremlin information warfare has evolved from human-driven troll farms to campaigns powered by generative artificial intelligence (AI) that are able to create deepfakes, fake online personas and synthetic news at unprecedented scale and speed.
- The war in Ukraine illustrates how generative AI enhances Russia's propaganda and psychological operations, acting as a test ground for tactics later employed elsewhere.
- Across Europe and the Global South, generative AI-driven manipulation increasingly threatens elections and public trust by blurring the line between fabricated and authentic information.
- Generative AI is not inherently a problem; its misuse is. The same technology used for manipulation can also build resilience. To protect information integrity, democracies should invest in generative AI-assisted detection and analysis tools, media literacy, and pre-bunking programs, and establish coordinated monitoring systems.

Introduction

The purposeful distortion of information has long been a political strategy but generative AI has accelerated its speed and scale (Kalpokas 2024). In contemporary Russia, it has become a fundamental component of foreign and security policy. Moscow views the information sphere not only as a supplement to traditional warfare but also as a battleground in its own right, where command of narratives can yield military, social and political advantages (Mitrović 2023). The Kremlin has consistently and adaptively invested in disinformation, influenced by both the lessons learned from previous campaigns and technology advancements (Polyakova and Fried 2019).

The significance of disinformation in hybrid warfare has been heightened by the war in Ukraine, which started in 2014 and sharply intensified in 2022. Observers have referred to it as "the first TikTok war" and "the first AI war," highlighting its role as a test ground for cutting-edge propaganda strategies and technological competition for informational supremacy (Chayka 2022). By utilizing information operations, Russia has attempted to demoralize Ukrainian soldiers and civilians on the battlefield

About the Author

Daniela Iampolsca is a researcher specializing in international relations, security, emerging technologies, strategic political communication, public diplomacy and multilateral governance. Her work focuses on the intersection of technology governance, strategic communication and global policy cooperation, with particular attention to their implications for peace, security and democratic resilience.

She previously served as an assistant project coordinator at the Hague Centre for Strategic Studies, where she supported the Global Commission on Responsible Artificial Intelligence in the Military Domain.

Daniela holds a master's degree in political science, with a specialization in public policy and governance, from the University of Amsterdam, and a bachelor's degree in international relations and organizations from Leiden University. Her academic and professional experience has provided her with a strong foundation in international security, governance frameworks and multilateral policy processes.

She is particularly interested in how strategic communication and technological innovation shape international security dynamics; influence policy making; and contribute to global peace, accountability and democratic resilience.

(Giles 2023). At the international level, Moscow uses disinformation as it tries to justify its aggression; undermine Western backing for Ukraine; and advance opposing narratives regarding the United States, the North Atlantic Treaty Organization (NATO) and the European Union (Martynyuk 2024).

Traditionally, the main sources of Russian disinformation have relied on troll farms, bot networks and the abuse of social media algorithms. Advances in generative AI have altered these dynamics, amplifying rather than replacing traditional techniques and serving as a force multiplier, boosting both volume and plausibility. Large-scale production of realistic text, images, audio and video by generative AI systems makes it possible to create fake narratives that are more convincing, harder to identify and simpler to distribute widely (Groumpos 2023; Endert 2024).

This policy brief examines how Moscow's AI-powered propaganda efforts, honed throughout the war in Ukraine, are changing the nature of information warfare, and suggests a multi-layered approach to combat this growing threat. It is mainly addressed to politicians in democratic governments, the European Union and NATO, as well as to technology companies, civil society and academic institutions that are crucial to information governance. The recommendations centre on enhancing international coordination, supporting media literacy and pre-bunking initiatives, and utilizing generative AI as a tool for monitoring and analysis rather than manipulation.

Russia's Changing Tool Kit: From Troll Farms to Transformer Models

Russia's disinformation tactics can be divided into three overlapping phases: troll farms, algorithmic exploitation and generative AI-powered campaigns. This is not because one phase totally supplanted the others, but rather because each

one represents a significant advancement in approach and capabilities, fuelled by strategic adaptability and technology advancement. While traditional strategies such as human-run online personas and fake news are still central, AI-generated material is enhancing them. Trolls and bots are not replaced by deepfakes or AI-generated personalities; they become more convincing, scalable and harder to identify.

Troll Farms: Human Manipulation as the Narrative Core

Human actors — including journalists at Kremlin-run media sites such as RT and Sputnik, agents at groups such as the Internet Research Agency and networks of like-minded influencers — are at the core of Russian disinformation (Paul and Matthews 2016). They generate narratives that are subsequently amplified by disinformation campaigns, portraying the West as hypocritical, Ukrainians as Nazis and NATO as an aggressor (Pomerantsev 2019). This human layer contributes strategic intent and contextual knowledge.

Algorithmic Exploitation: The Multipliers

Automation increases reach. Fake accounts, bot networks and algorithm manipulation flood the information environment, push hashtags into trending topics and provide the illusion of consensus (Woolley and Howard 2018). These instruments ensure that Kremlin-supported narratives seem to be everywhere, influencing coverage in the mainstream media. Although platforms have improved detection, Russia still uses advanced bot networks that imitate real people (DiResta et al. 2019). These bots serve as multipliers when combined with narratives that have been hand-picked by humans.

Generative AI Enhancement: The Force Multiplier

Generative AI represents a revolutionary development in Russian information operations. In its 2019 national strategy for AI development, Russia pledged to become a global leader in AI by 2030 (Hybrid Warfare Analytical Group 2025). Generative AI offers a force multiplier: the capacity to produce convincing, scalable and adaptable content at volumes and rates that greatly surpass those of conventional techniques (Endert 2024).

The production of synthetic content is the most visible use of generative AI in disinformation. Nowadays, large language models can produce convincing blog posts, news articles and social media comments at scale, customized for various audiences and contexts (Barman, Guo and Conlan 2024). Unlike earlier bot networks that frequently exposed themselves through poor grammar or clumsy repetition, AI-generated writing can shift tone, vocabulary and arguments in ways that resemble real-world human communication (Brandizzi 2023). This makes automated accounts seem more legitimate and harder for fact-checkers to detect.

Ukraine as the Front Line of Information Warfare

Ukraine serves as a stark example of how Russia's multi-layered disinformation ecosystem operates in reality. Moscow's tactics have been heavily reliant on disinformation since 2014, when fake reports of crimes and propaganda about "fascists in Kyiv" were used to justify the annexation of Crimea (Meister 2016). Since 2022, these operations have intensified and increasingly use generative AI to make progress in Russia's full-scale invasion of Ukraine.

The propaganda used against Ukraine simultaneously:

- → specifically targets Ukrainians in an effort to lower morale;
- → aims to weaken public support for Kyiv in the West; and
- → gives a preview of strategies that might be applied in elections and crises elsewhere.

Ukraine serves as an early warning system for democracies. The techniques utilized there could be adapted against NATO and EU states, demonstrating not only the speed but also the flexibility of generative AI-powered disinformation (Padalko 2025). Campaigns can appear within hours of actual occurrences, shaping the early information environment before reliable reporting can react (Chesney and Citron 2019). This accelerated pace makes it harder to maintain a

stable information environment and increases the risk of confusion during emergencies.

The creation of deepfakes is a striking example of generative AI's power in information warfare. These artificially created audio and video forgeries have already been used by Russia during the war in Ukraine, most notably in March 2022, when a fake video purportedly showing Ukrainian President Volodymyr Zelensky calling for the surrender of Ukrainian troops went viral (Allyn 2022; Pearson and Zinets 2022). Although quickly debunked, this example illustrates that even crude deepfakes can briefly cause confusion and force governments, media and platforms to invest time and resources into verification and public reassurance.

As technology advances, this type of video can become increasingly indistinguishable from real footage, misleading viewers and raising concerns about the reliability of authentic recordings (Singh and Dhumane 2025). By making sure that any piece of information can be questioned, even those that are genuine, deepfakes undermine what academics refer to as epistemic vigilance, or a society's capacity to distinguish between truth and lies (Sperber et al. 2010).

In addition to affecting Ukrainians, Russian disinformation also targets foreign audiences. In an effort to influence Western voters and put pressure on democratic governments to reduce support for Ukraine, Kremlin-affiliated networks push narratives portraying Ukraine as corrupt, unappreciative or incapable of winning the war. Russia's use of an AI-powered tool called Meliorator, which generated more than 1,000 fake American social media profiles to push anti-Ukraine, pro-Kremlin propaganda, is a startling example of how the same ecosystem that targets Ukrainians is also designed to influence Western political debates (Harding 2024).

The ability of generative AI to personalize and microtarget may be its most strategically important use. By combining disinformation with generative AI-driven analytics, Russia can tailor content to exploit specific political and cultural vulnerabilities, with generative AI enhancing the potential reach, speed and adaptability of campaigns across numerous audiences simultaneously (Padalko 2025). Recent examples show how these strategies are refined in Ukraine and further disseminated elsewhere. AI-generated accounts have propagated inaccurate allegations of criminal activity by Ukrainian refugees in Europe in an effort to inflame xenophobia and erode support for Kyiv (Morris and Oremus 2022).

In November 2025, the Center for Countering Disinformation in Ukraine reported that a fresh batch of AI-generated fake videos purporting to show a "mass surrender" of Ukrainian forces in the vicinity of Pokrovsk was circulating on TikTok (Kohanets 2024). The videos, which were promoted by coordinated accounts and distributed in several languages, aimed to demoralize Ukrainians and persuade foreign viewers that Ukraine was losing the war and that additional aid was pointless. This reflects the dual-purpose strategy of Russian information operations: test and target domestically, then expand internationally.

Global Implications of Generative AI-Driven Disinformation

The dynamics seen in Ukraine are not isolated; they provide insight into how Russian generative AI-powered operations can expand internationally. Generative AI's application in disinformation operations has revolutionized the way governments and non-state entities may affect societies worldwide (Chesney and Citron 2019). Numerous strategies that Russia initially employed or refined in Ukraine — such as narrative manipulation, generative AI-assisted amplification networks and rapid content creation — have since emerged in other geopolitical contexts. Generative AI-driven disinformation can now be rapidly modified and disseminated internationally, making it no longer limited to the immediate theatre of war (Bachmann, Putter and Duczynski 2023).

Election interference is a major concern. Russian propaganda has long been targeting elections around the world, including in the United States, but generative AI has given these campaigns new impetus (De Luce and Collier 2024). Generative AI-enabled networks can generate and disseminate false information on a large scale, increasing informational noise and making it harder to discern trustworthy reporting. A notable example is the "DoppelGänger" campaign, in which AI-generated "news sites" imitate the appearance

of authentic Western outlets to promote Kremlin narratives (Antoniuk 2023). Russia's foreign meddling now follows the same approach it has employed in Ukraine: flood the area, distort the narrative space and overwhelm verification.

This global diffusion is starkly illustrated by a recent example generated before Poland's presidential election. According to a BBC Verify investigation, a network with ties to Russia used AI to mimic the voices of British emergency personnel (Robinson 2025). An emergency medical adviser from England, who was one of the victims, was horrified to learn that a disinformation campaign had used a synthetic version of his voice (ibid.). Using fabricated statements about impending terrorist strikes that were purportedly scheduled to coincide with the presidential election, the cloned voice was inserted into a video, lending credibility to misleading claims. The operation is similar to previous Russian initiatives in Ukraine to use fake political speeches, showing how front-line experimentation facilitates further global deployment (Allyn 2022; Pearson and Zinets 2022).

This incident highlights two crucial aspects of disinformation enhanced by generative AI. First, it illustrates the increasing complexity of synthetic media, which now includes extremely life-like voice cloning in addition to text and video. Second, it demonstrates the global reach of such campaigns, with voices from the United Kingdom incorporated in content disseminated during the Polish election. Such activities demonstrate how Russia uses generative AI to directly interfere in Western political systems in addition to using it to sway narratives at home and on the battlefield.

These threats are not limited to Western democracies. Russia is increasingly deploying deception in the Global South, using deepfakes, avatar networks and AI-generated content to portray itself as a champion of anti-colonial resistance (Zhang, Jin and Si 2025). These initiatives involve state media, allied organizations and local proxies in states such as Burkina Faso, the Central African Republic and Mali (Ehl and Ghaedi 2025; Nilsson-Julien 2025). Russia simultaneously presents itself as a liberator, victim and defender of sovereignty, all while waging imperial-style wars. Many efforts prioritize framing and interpretation above factual claims, utilizing historical grievances and anti-colonial narratives (Digital Forensic Research Lab 2024). This tactic turns post-colonial solidarity into a tool for soft power and geopolitical influence at a minimal cost.

These dynamics show that Ukraine serves as an early indicator of how generative AI-enabled disinformation strategies may be adopted elsewhere. There is growing evidence that the strategies developed in Ukraine — such as coordinated inauthentic networks, AI-generated personas and fake news websites — are already being used in North America and Europe (Nimmo and Agranovich 2022; Antoniuk 2023). The same disinformation ecosystem that undermines Western support for Kyiv is now adaptable to new political contexts and languages. Generative AI-enhanced disinformation must thus be viewed as a global challenge rather than as a regional issue exclusive to Ukraine.

Policy Recommendations

Combatting generative AI-enabled disinformation requires transnational, coordinated and persistent action. The following recommendations provide a framework for strengthening democratic resilience in the era of propaganda driven by generative AI.

Create AI Disinformation Watch Centres to Ensure a Coordinated International Response

The first task is to enhance collective awareness. The European Union and NATO should establish AI disinformation watch centres to monitor, evaluate and disseminate intelligence on new disinformation strategies. In NATO's case, this could build upon already-existing frameworks, such as the Information Environment Assessment (NATO Allied Command Transformation 2019), expanding them to systematically address AI-driven narrative manipulation and synthetic media. To provide early warning of new types of synthetic media or disinformation efforts, these centres would serve as permanent hubs connecting technological businesses, university researchers and intelligence services. By combining real-time data and technical know-how, these centres could help identify emerging patterns at an early stage.

They would also improve military-civilian coordination, bridging the information policy

and cyber defence divide. Beyond Europe, these centres should work closely with allies abroad to create common threat assessments. Joint response mechanisms and rapid intelligence sharing would make it more difficult for state and non-state actors, including Russia, to exploit jurisdictional or platform regulatory gaps.

Invest in Media Literacy, Critical Thinking and Pre-bunking

Although technological solutions are vital, it is becoming more widely acknowledged that a society's cognitive resilience is a crucial line of defence against generative AI-assisted information operations (Gerlich 2025). As demonstrated in previous sections, Russia's generative AI-enabled efforts focus not only on fake visuals but also on identity-based framing, emotionally charged narratives and historical grievances. Generative AI frequently targets cognitive biases rather than factual gaps, as evidenced by the Pokrovsk deepfake campaign and the extensive dissemination of AI-generated refugee-crime narratives. These strategies cannot be defeated by technical detection alone because they function at the interpretative rather than the factual level.

Pre-bunking campaigns, or initiatives that alert citizens to deceptive tactics before they come across them in the wild, should receive more funding from governments, civil society organizations and educational institutions. Practical models already exist. For instance, the Taiwan FactCheck Center combines rapid fact-checking with public explanations of manipulation strategies, helping audiences identify deceptive narratives rather than merely respond to individual false claims (Chen 2025). According to Stephan Lewandowsky and Sander van der Linden (2021, 10), pre-bunking, like a vaccination, develops "mental antibodies" against disinformation by educating people to identify emotional triggers, logical fallacies and incorrect framing. However, to be effective, these initiatives should prioritize narrative framing, agenda setting and source evaluation instead of visual authenticity, which has become less dependable with generative AI (Nightingale and Farid 2022).

Special attention should be given to audiences shown to be more susceptible to false narratives, such as elderly users and populations with low levels of confidence in traditional media (Guess, Nagler and Tucker 2019). Programs should also be linguistically and culturally regionalized, particularly where Russian campaigns make use of local historical or post-colonial narratives, such as parts of the Global South (Ehl and Ghaedi 2025; Nilsson-Julien 2025).

Utilize Generative AI as a Support Tool for Detection and Analysis

Although generative AI is frequently associated with producing propaganda, it can also support democratic responses to disinformation by assisting with detection, monitoring and analysis. With human oversight, generative AI can help institutions identify coordinated behaviour and track narrative diffusion across platforms, though it cannot fully stop or neutralize disinformation on its own (Starbird, Arif and Wilson 2019).

Governments and tech firms should selectively invest in generative AI-assisted tools that improve analytical capacity, such as systems for tracking the spread of multilingual narrative propagation, detecting coordinated inauthentic behaviour and assisting with provenance assessment for synthetic media. These instruments should supplement human judgment and operate within explicit accountability and transparency guidelines.

However, policy makers should also be aware of the limitations of generative AI-based defences. Detection does not automatically result in correction, and automated labelling or takedowns may have inconsistent or unexpected consequences on public trust (Chesney and Citron 2019; Nightingale and Farid 2022). Generative AI-enabled monitoring should be integrated into broader institutional frameworks, including public education, collaboration with independent media, crisis communication and legal protections, making it one part of a multi-layered, governance-driven response.

Conclusion

The use of generative AI as a tool in disinformation strategies marks a turning point in the history of information warfare. Once reliant on human trolls, it has evolved into a decentralized, automated ecosystem that can instantly reach and influence global audiences. The

ramifications go far beyond Ukraine's borders: generative AI-powered propaganda threatens democracy by blurring the truth, eroding public trust and spreading misleading narratives.

Building resilience requires updated strategies that combine international coordination, civic education, technological innovation and adaptive regulation. To ensure that generative AI becomes a pillar of information integrity rather than a tool for enhancing deception, governments, tech firms and civil society organizations must work in close coordination. Generative AI-assisted detection, monitoring and pre-bunking should complement human oversight within multi-layered governance frameworks.

Preserving the integrity of public discourse in the digital age requires more than just reactive fact-checking. Information operations frequently depend more on interpretation, framing and conflicting narratives shaped by power and history, instead of objective facts. Addressing these challenges necessitates foresight, adaptability and coordination. Democracies can reduce their vulnerability by enhancing the conditions for informed debate and by investing in generative AI-aware defensive tools and digital literacy programs that assist citizens in navigating complicated information environments.

Works Cited

Allyn, Bobby. 2022. "Deepfake video of Zelenskyy could be 'tip of the iceberg' in info war, experts warn." NPR, March 16. www.npr.org/2022/03/16/1087062648/deepfake-video-zelenskyy-experts-war-manipulation-ukraine-russia.

Antoniuk, Daryna. 2023. "Russia-linked 'Doppelgänger' social media operation rolls on, report says." The Record, December 5. https://therecord.media/doppelganger-influence-operation-new-activity.

Bachmann, Sascha-Dominik Dov, Dries Putter and Guy Duczynski. 2023. "Hybrid warfare and disinformation: A Ukraine war perspective." *Global Policy* 14 (5): 858–69. https://doi.org/10.1111/1758-5899.13257.

Barman, Dipto, Ziyi Guo and Owen Conlan. 2024. "The Dark Side of Language Models: Exploring the Potential of LLMs in Multimedia Disinformation Generation and Dissemination." *Machine Learning with Applications* 16: 100545. https://doi.org/10.1016/j.mlwa.2024.100545.

Brandizzi, Nicolo'. 2023. "Toward More Human-Like AI Communication: A Review of Emergent Communication Research." *IEEE Access* 11: 142317–40. https://doi.org/10.1109/ACCESS.2023.3339656.

Chayka, Kyle. 2022. "Watching the World's First 'Tik Tok War.'" *The New Yorker*, March 3. www.newyorker.com/culture/infinite-scroll/watching-the-worlds-first-tiktok-war.

Chen, Summer. 2025. *Countering AI Disinformation: Lessons from Taiwan's 2024 Election Defense Strategies.* Research paper, Information Resilience & Integrity Symposium. https://saferinternetlab.org/wp-content/uploads/2025/08/Panel-4-Summer.pdf.

Chesney, Robert and Danielle K. Citron. 2019. "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security." *California Law Review* 107: 1753–819. https://doi.org/10.15779/Z38RV0D15J.

De Luce, Dan and Kevin Collier. 2024. "Russia's 2024 election interference has already begun." NBC News, February 26. www.nbcnews.com/news/investigations/russias-2024-election-interference-already-begun-rcna134204.

Digital Forensic Research Lab. 2024. "Two-pronged approach to Africa pays dividends for Russia." Atlantic Council, February 29. www.atlanticcouncil.org/in-depth-research-reports/issue-brief/two-pronged-approach-to-africa-pays-dividends-for-russia/.

DiResta, Renée, Kris Shaffer, Becky Ruppel, David Sullivan, Robert Matney, Ryan Fox, Jonathan Albright and Ben Johnson. 2019. "The Tactics & Tropes of the Internet Research Agency." US Senate Document. October. https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1003&context=senatedocs.

Ehl, David and Monir Ghaedi. 2025. "Russian propaganda: How Moscow uses disinformation in Africa." Deutsche Welle, February 3. www.dw.com/en/how-russias-propaganda-machine-sows-disinformation-in-africa/a-71453082.

Endert, Julius. 2024. "Generative AI is the ultimate disinformation amplifier." Deutsche Welle, March 17. https://akademie.dw.com/en/generative-ai-is-the-ultimate-disinformation-amplifier/a-68593890.

Gerlich, Michael. 2025. "AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking." *Societies* 15 (1). https://doi.org/10.3390/soc15010006.

Giles, Keir. 2023. "Russian cyber and information warfare in practice." Research paper, December 14. Chatham House. www.chathamhouse.org/2023/12/russian-cyber-and-information-warfare-practice.

Groumpos, Peter P. 2023. "A Critical Historic Overview of Artificial Intelligence: Issues, Challenges, Opportunities, and Threats." *Artificial Intelligence and Application* 1 (4): 181–97. https://doi.org/10.47852/bonviewAIA3202689.

Guess, Andrew, Jonathan Nagler and Joshua Tucker. 2019. "Less than you think: Prevalence and predictors of fake news dissemination on Facebook." *Science Advances* 5 (1). https://doi.org/10.1126/sciadv.aau4586.

Harding, Emily. 2024. "A Russian Bot Farm Used AI to Lie to Americans. What Now?" Center for Strategic & International Studies, July 16. www.csis.org/analysis/russian-bot-farm-used-ai-lie-americans-what-now.

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

Kalpokas, Ignas. 2024. "Post-Truth and Information Warfare in their Technological Context." *Applied Cybersecurity & Internet Governance* 3 (2) 99–121. https://doi.org/10.60097/ACIG/190407.

Kohanets, Roman. 2025. "Deepfake Blitz: Kremlin Pushes AI Videos of Ukrainian Troops 'Giving Up' Near Pokrovsk." United24 Media, November 5. https://united24media.com/latest-news/deepfake-blitz-kremlin-pushes-ai-videos-of-ukrainian-troops-giving-up-near-pokrovsk-13114.

Lewandowsky, Stephan and Sander van der Linden. 2021. "Countering Misinformation and Fake News Through Inoculation and Prebunking." *European Review of Social Psychology* 32 (2): 348–84. https://doi.org/10.1080/10463283.2021.1876983.

Martynyuk, Leonid. 2024. "Russian propaganda entangles NATO, West, and Ukraine in knot of 'dangerous,' 'provocative' rhetoric." Voice of America, March 24. www.voanews.com/a/7625668.html.

Meister, Stefan. 2016. "Isolation and Propaganda: The Roots and Instruments of Russia's Disinformation Campaign." Transatlantic Academy Paper Series No. 6. April. https://dgap.org/system/files/article_pdfs/meister_isolationpropoganda_apr16_web_1.pdf.

Mitrović, Miroslav. 2023. "Russian Strategic Communication – Endless Information Warfare." *Security Science Journal* 3 (2). https://doi.org/10.37458/ssj.3.2.2.

Morris, Loveday and Will Oremus. 2022. "Russian disinformation is demonizing Ukrainian refugees." *The Washington Post,* December 8. www.washingtonpost.com/technology/2022/12/08/russian-disinfo-ukrainian-refugees-germany/.

NATO Allied Command Transformation. 2019. "Fact Sheet – Information Environment Assessment (IEA)." May. Norfolk, VA: NATO Allied Command Transformation. www.act.nato.int/wp-content/uploads/2023/05/2019_05_IEA.pdf.

Nightingale, Sophie J. and Hany Farid. 2022. "AI-synthesized faces are indistinguishable from real faces and more trustworthy." *Proceedings of the National Academy of Sciences* 119 (8). https://doi.org/10.1073/pnas.2120481119.

Nilsson-Julien, Estelle. 2025. "The African Initiative: Russian-backed outlet peddles propaganda to buy influence in Africa." Euronews, June 18. www.euronews.com/my-europe/2025/06/18/the-african-initiative-russian-backed-outlet-peddles-propaganda-to-buy-influence-in-africa.

Nimmo, Ben and David Agranovich. 2022. "Removing Coordinated Inauthentic Behavior From China and Russia." Meta, September 27. https://about.fb.com/news/2022/09/removing-coordinated-inauthentic-behavior-from-china-and-russia/.

Padalko, Halyna. 2025. "AI and Information Manipulation: Russia's Interference in the US Elections." Digital Policy Hub Working Paper. www.cigionline.org/static/documents/DPH-paper-Padalko_7Fz4cUS.pdf.

Paul, Christopher and Miriam Matthews. 2016. "The Russian 'Firehose of Falsehood' Propaganda Model." Expert Insights, July 11. Santa Monica, CA: RAND Corporation. www.rand.org/pubs/perspectives/PE198.html.

Pearson, James and Natalia Zinets. 2022. "Deepfake footage purports to show Ukrainian president capitulating." Reuters, March 16. www.reuters.com/world/europe/deepfake-footage-purports-show-ukrainian-president-capitulating-2022-03-16/.

Polyakova, Alina and Daniel Fried. 2019. *Democratic Defense Against Disinformation* 2.0. Report, Atlantic Council. June. www.atlanticcouncil.org/wp-content/uploads/2019/06/Democratic_Defense_Against_Disinformation_2.0.pdf.

Pomerantsev, Peter. 2019. *This Is Not Propaganda: Adventures in the War Against Reality*. London, UK: Faber & Faber.

Robinson, Olga. 2025. "British 999 call handler's voice cloned by Russian network using AI." BBC, July 31. www.bbc.com/news/videos/c3dpeyrx1kyo.

Singh, Sonam and Amol Dhumane. 2025. "Unmasking digital deceptions: An integrative review of deepfake detection, multimedia forensics, and cybersecurity challenges." *MethodsX* 15: 103632. https://doi.org/10.1016/j.mex.2025.103632.

Sperber, Dan, Fabrice Clément, Christophe Heintz, Olivier Mascaro, Hugo Mercier, Gloria Origgi and Deirdre Wilson. 2010. "Epistemic Vigilance." *Mind & Language* 25 (4): 359–93. https://doi.org/10.1111/j.1468-0017.2010.01394.x.

Starbird, Kate, Ahmer Arif and Tom Wilson. 2019. "Disinformation as Collaborative Work: Surfacing the Participatory Nature of Strategic Information Operations." *Proceedings of the ACM on Human-Computer Interaction* 3, 127: 1–26. https://doi.org/10.1145/3359229.

Woolley, Samuel C. and Philip N. Howard, eds. 2018. *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media*. Oxford, UK: Oxford University Press.

Zhang, Chang, Yong Jin and Ran Si. 2025. "Propaganda à la Russe: historical continuance and modern adaptation." *Critical Studies in Media Communication* 42 (1): 69–74. https://doi.org/10.1080/15295036.2025.2469094.

About CIGI

The Centre for International Governance Innovation (CIGI) is an independent, non-partisan think tank whose peer-reviewed research and trusted analysis influence policy makers to innovate. Our global network of multidisciplinary researchers and strategic partnerships provide policy solutions for the digital era with one goal: to improve people's lives everywhere. Headquartered in Waterloo, Canada, CIGI has received support from the Government of Canada, the Government of Ontario and founder Jim Balsillie.

À propos du CIGI

Le Centre pour l'innovation dans la gouvernance internationale (CIGI) est un groupe de réflexion indépendant et non partisan dont les recherches évaluées par des pairs et les analyses fiables incitent les décideurs à innover. Grâce à son réseau mondial de chercheurs pluridisciplinaires et de partenariats stratégiques, le CIGI offre des solutions politiques adaptées à l'ère numérique dans le seul but d'améliorer la vie des gens du monde entier. Le CIGI, dont le siège se trouve à Waterloo, au Canada, bénéficie du soutien du gouvernement du Canada, du gouvernement de l'Ontario et de son fondateur, Jim Balsillie.

Credits

Research Director, Digitalization, Security & Democracy Aaron Shull
Program Manager Reanne Cayenne
Publications Editor Christine Robertson
Graphic Designer Sepideh Shomali

67 Erb Street West
Waterloo, ON, Canada N2L 6C2
www.cigionline.org

CIGI