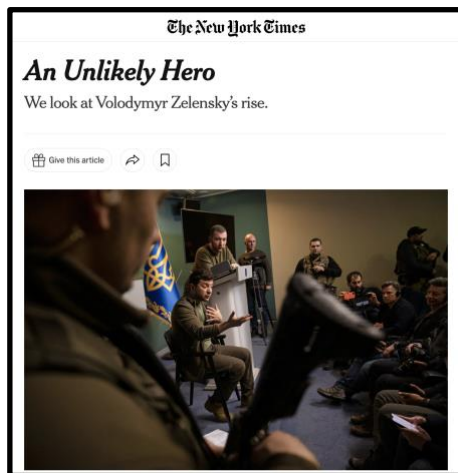


A Faceless Enemy: AI Propaganda in the Russo-Ukrainian War

Will Forsen

Introduction

The invasion and ongoing conflict in Ukraine were initiated by Putin and Russian troops on February 24, 2022. This is not a new conflict: Ukraine has undergone waves of pro-Russian separatist groups in the eastern Donbas region since 2014 after Russia's illegal annexation of Crimea. Purporting to protect their Russian-speaking population and historic land, Russia has since increased media disinformation as well as military aggression. News of mid-April conveys Russia's decreased focus around the capital of Kyiv as they prepare for renewed fighting in the east of Ukraine. Ukrainian President Zelenskyy's brave struggle for democracy in the face of violence is to be commended on no uncertain terms. Or, that is what you may believe if you keep up with the New York Times.¹ A Russian citizen or sympathizer, however, may see him as "the man who sold Ukraine," "relentlessly provoking Russia[.]"² Which do you believe, and why?



Web 2.0 is the current, social-networking version of the Internet in which content is generated by users, as opposed to the earlier form of published static information. As the Internet grew into the loud platform as we know it, it became all too possible that consumers digest a

manipulated version of the truth. In the case of the current war in Ukraine, such manipulation is not only likely but guaranteed by the high stakes various parties possess in the conflict. This propaganda takes a different form than stereotypes and movies would have the reader envision: gone are the days of Uncle Sam pointing and hastily drawn caricatures of Cold War communists. Communication outlets across the world now engage in a new, nefarious crossover of fabrications with artificial intelligence to increase the scope and intensity of their messaging. The truly disturbing aspect of this messaging? It's working.

Russian news outlets are not well-known for their objectivity in reporting—in fact, the Press Freedom Index ranks the country at 142 out of 179 countries.³ Anyone following the news reported on Western sites has encountered articles describing some of the Kremlin's more far-fetched headlines: "Ukrainian-American Biolabs' Existence Proved by Various Sources"; Real Danger of Western Pfizer and Moderna Vaccines"; and "Denazification of Ukraine: Operational Summary."⁴ Two of the most popular sites for such propaganda are NewsFront and SouthFront, outlets which espouse favorable narratives for the Russian government. Although these narratives are pushed by conventional websites and articles, questions about the involvement of AI in such propaganda have begun circulating.

Have NewsFront and SouthFront implemented AI deep fakes and web manipulation in their presentation of the Russo-Ukrainian War? What implications do such applications signify in our current Age of Information? This case study will examine the role of these sites and their propaganda in both domestic Russian discourse as well as foreign. While separate studies on propaganda's effects on Russian or Western opinions may provide better focus, there, unfortunately, remains too little accessible evidence and news in the former case to produce a clear enough picture. A substantial study on Russian propaganda's influence on Russians is elusive, and

the nature of this complication is discussed below. Furthermore, the effects of Russian-sourced propaganda are felt in both domestic and foreign spheres with similar thought-provoking implications. For these reasons, I attempt to broach the subject of both domestic and abroad reactions to AI-associated political propaganda. The overt ethical implications of propaganda and truth naturally take a prominent role in such a discussion. There also exist several obstacles for those who utilize AI for widespread communication. To approach these topics and more, however, it is useful to begin by determining AI's historic and current role in both communication and politics. From there we may judge misinformative efforts and further questions.

AI in politics and communication

Propaganda is a popular and historic way of conveying information to groups, usually national populaces, in a way that affirms or seeks to establish an official position. It purports to signify truth-value while frequently presenting opaque, one-sided, or patently false statements. Russia and the U.S., and even your mom on Facebook, are all content creators susceptible to propaganda's effects just as they are capable of producing it.

Information-sharing then becomes a controversial field, and groups seek to capitalize on possible misinformation to promote their position or at least detract from their opponents. AI has begun to factor into this contest for truth and influence. Programmers will now write algorithms to automate messages, manipulate visuals, and guide web users' online experience. A recent Carnegie Mellon study saw that about half of the accounts tweeting about the COVID-19 pandemic were bots, or an automated software application designed to resemble human activity online.⁵ Deep fakes are named after the deep learning form of AI and are designed to simulate a real person in videos, photographs, or audio.⁶ Not inherently malicious, deep fakes gather copious data to usually swap preexisting images, videos, etc. with their desired person. The more troublesome deep fakes

are users creating nonexistent speeches by world leaders, unfortunately not Reddit users imagining a world in which Keanu Reeves took the blue pill instead of the red.⁷

AI has amplified the potential influence of propaganda through these technologies and their convincing messaging. With the explosion of the Internet and sites such as Facebook and Twitter, it becomes increasingly more difficult to monitor and regulate how information is conveyed between parties. The rapid nature of most consumers' interactions with (social) media sites also betrays their quest for objectivity: bots are relatively "weak" AI, but can convince someone who takes only a few seconds to read a seemingly innocuous tweet. Algorithm programmers then influence their target audience in several ways: not only with superficially convincing material but also with web manipulation. This can take several of the following forms, and more: "The exploitation of human biases detected by AI algorithms, personalized addictive strategies for consumption of (online) goods, or taking advantage of the emotionally vulnerable state of individuals to promote products and services that match well with their temporary emotions."⁸ The last example, which plays on users' emotions, is particularly important to note for the current situation in Ukraine, as a tense and divisive topic across the world. Applied correctly, AI can influence web users' social media feeds to present them with biased news that confirms their perspective on the issue. The user's perspective can be derived from the personal information they provide when they register with a social site, as well as from their "likes" and general interaction with the posts they view. In the Carnegie Mellon example above, researchers observed that the bots' intentions were to "[sow] division in America" by spreading conspiracy theories.⁹ The implications for such technology and its ability to sway unknowing media consumers are vast and disturbing. The following section on ethics discusses ways in which to judge technology and consumers' habits.

Implications of usage

The ethical implications of AI in propaganda are numerous and overt. To begin, the medium through which AI works is already questionable—although not intrinsically negative, propaganda certainly carries that connotation, and its applications are usually widely condemned. Attaching AI to this tool of (political) persuasion simply augments the message's power. It is then more productive to discuss the ethics of propaganda and any distinctions that are made with the introduction of AI.

Propaganda is most dangerous when delivered in a vacuum of competition. In most of the Western world, individual consumers are privileged with a plethora of news outlets from which to digest current events, from Ukraine to the Kardashians. In parts of the world, however, free press from multiple sources is not guaranteed. NewsFront and SouthFront convey Putin's party lines while Western media is restricted, even more so after the initial February invasion.¹⁰ Without



opposing sources to counter the false narratives presented to them, most Russian citizens only have access to official state sources which rarely mention the war, if at all.¹¹ This is problematic because consumers cannot refute the information on any knowledgeable ground. They also cannot provide or support an alternative narrative due to the lack of source options. This effectively weakens both the (political) opposition and the credibility of any external media that slips in because there is little chance to corroborate information. Along a similar line, unaware Russian consumers have less of a reason to believe the occasional snippet of Western news because most of their own, homegrown information

contradicts it. AI bots and deep fakes then gain their *human* support as Russian citizens access global social media and repeat the narratives that have indoctrinated them. This legitimizes accounts that may otherwise seem artificial and as well as perspectives that run completely counter to Western media viewpoints, sowing division and widening the gap of understanding.

The practical implication, namely political division, is a concern for policymakers and the free press the world over. It becomes clear that AI accounts do not need to convince Western minds of Russia's justification in the war, they need only to spread uncertainty about the conflict itself. This is a victory in its way, as it complicates Western policymakers' relations with their public, the latter of whom may now be more apprehensive of encouraging sanctions and other actions taken against Russia. Without a united base to lean on for support, policymakers' decisions are hindered. The same decisions are also likely to have less stamina in support and overall media coverage of their target and effect. Not every global event is reported equally, and those that hold the public eye (i.e. headliners) receive larger donations for aid, presumably also symbolizing moral support for the recipient and policies taken to support them.¹² Frequently, however, such events lose momentum and popularity in the press. The media coverage of the Russo-Ukrainian War has already experienced a type of "burnout"; while the crisis may have dominated the cover of the New York Times and Wall Street Journal through late February and most of March 2022, holding public attention is difficult. Medical experts have compared increasing disinterest in the Ukraine conflict too much of the world's current aversion to any headline regarding COVID-19.¹³ AI-assisted misinformation encourages this disinterest by complicating research into the subject—average Internet users have a "shorter attention span than a goldfish," and conflicting information does not fit into the desire for instant satisfaction of verified news.¹⁴

AI expands the range of political propaganda through its web manipulation and attempts at human messaging, resulting in ethical and practical issues exacerbated beyond manual reach. However successful this section may portray the nefarious efforts of propagandists, there remain several obstacles for the deployment of convincing AI communications.

Obstacles for users

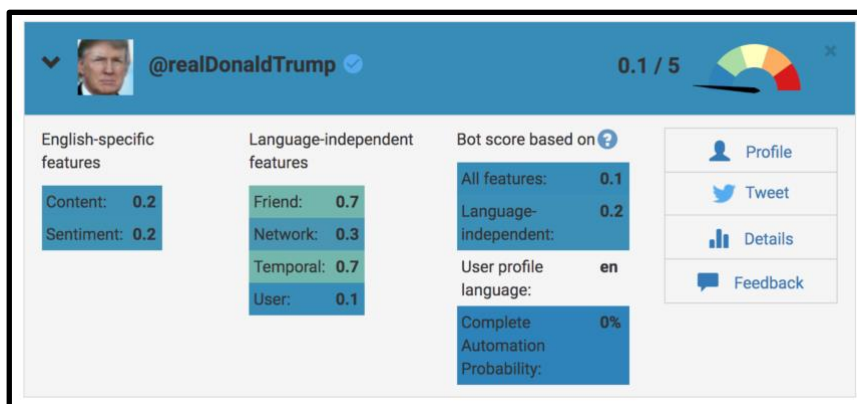
Practical obstacles to AI-assisted propaganda include convincing the masses and the AI technology itself. The first point is self-explanatory: to set up a narrative or place an existing one into question, AI bots must appear authentic enough to deceive users into believing the message or even reposting it. Although the range of social media is amplified by rapid AI bots, this means that the same message reaches people with diverse experiences and knowledge. As a particular fabrication becomes popular on the Internet, it is increasingly likely that someone will notice and raise the alarm, or at least begin a debate that encourages critical thinking.

Competing and troublesome technologies seem to be the most potent threats AI propaganda faces. A Brookings Institution article neatly outlines two of the primary obstacles for those wishing to spread disinformation: bot detection software and the sheer volume of necessary data for machine learning algorithms.¹⁵ Programmers seek to legitimize accounts in various ways to avoid these obstacles. This usually involves setting up bots before posting to gather enough data to have intelligent interactions. Bots may follow each other and real users to further normalize online activity.

Even as bots collaborate to post similar fake news, more AI applications are popping up to judge their veracity. AI bot detection is being countered by tools such as one from the University of Waterloo, specifically used to “determine if claims made in posts or stories are supported by

other posts and stories on the same subject.”¹⁶ Not meant to operate autonomously, this tool and others increase the capacity and speed of an Internet consumer to fact-check

Indiana University’s Botometer ranks the bot-like behavior of Twitter accounts. When one inputs Donald Trump, he receives a result of 0% of “complete automation probability.”



questionable sources. Individuals then remain the most formidable barriers to propagandists, just as they are also the key targets. Dishonest news outlets and programmers must design convincing codes to trick consumers and keep up with the advancing technology of watchdogs. Now we may turn to the specific example of Russia and two of the country’s media sources as to their role in the Ukraine conflict.

Russian media in the context

Russia’s media landscape is unique in its geographic, linguistic, and cultural diversity. News outlets have adapted to a country that spans eleven time zones and contains hundreds of languages and dialects, and most citizens consume TV media. Top networks are either state-run or have close ties with the Kremlin, and this is mirrored in other forms of media even as the Internet becomes a more popular source of information. As it has grown in popularity, however, authorities have expanded powers that suppress or restrict information online, passing laws to allow the

blocking of certain content and to render more user data accessible. The 2019 Sovereign Internet Law granted the government the power to effectively disconnect from the global Internet by routing web traffic through Kremlin-operated infrastructure.¹⁷ The implications of such an “online Iron Curtain” are enormous: regardless of feasibility questions, this calls for the objectivity of every piece of information available to Russian citizens.¹⁸

Although Russia has not had the most illustrious history with a free press, recent amendments to laws on media information dispelled all doubts as to Putin’s desire for misinformation and secrecy. The March 4th amendments to the criminal code include fines and jail terms for up to 15 years for those who distribute “fake” information.¹⁹ As a result, a myriad of respected outlets such as BBC, CBS, and Bloomberg pulled reporters and halted broadcasting.²⁰ The motivation for enacting such legislation is obvious: to control the content and stream of information to the Russian people. Two outlets mentioned earlier, NewsFront and SouthFront, seem to be accomplishing just that goal with their targeted and skewed presentation of the conflict in Ukraine.

Both SouthFront and NewsFront were founded in 2015 with the ostensible goal of “providing an ‘alternative source of information’ for Western audiences.”²¹ Both are Crimea-based sources with extensive ties and funding relationships with the Kremlin, including intelligence agencies.²² The U.S. Department of the Treasury has deemed the sites as direct arms of the Russian government, and others have made connections between them and Twitter and Facebook profiles espousing misinformation regarding the conflict.^{23,24} While some attempts of deception are somewhat obvious, such as a deep fake video of President Zelenskyy calling for Ukrainians to lay

down their arms, the influence of similar efforts should not be discounted. There is difficulty in quantifying the influence of AI accounts and bot messages besides online “likes” and retweets, but News- and SouthFront have succeeded in at least sowing uncertainty abroad. They are also privileged to present the one narrative available to the Russian people at home, thanks to the aforementioned media amendments.

A senior reporter at NBC News, Ben Collins, discusses on his Twitter the extensive efforts Russian AI propaganda has gone to in



misleading not only their citizens but Ukrainians as well.²⁵ News- and SouthFront have garnered attention in this way, but perhaps more notably from the U.S. State Department in 2016 for their disputed U.S. presidential election interference. Facebook removed accounts associated with SouthFront in 2020 after a spokesperson said a propaganda campaign had used computer-generated faces to bolster the credibility of fake columnists across several platforms, including Instagram.²⁶ A special report by the State Department in the same year outlines the history and context of this site and NewsFront which led to such a situation. The report, titled “Pillars of Russia’s Disinformation and Propaganda Ecosystem,” however, does not mention the words “artificial intelligence” or “bots.”²⁷ A lot has happened since 2020, especially at this point in mid-April of 2022, and governments and social media sites are becoming increasingly aware of such automated attempts to distort public opinion. Although Putin’s initial hopes for a swift invasion and clear victory have faded, the Kremlin seems as determined as ever to utilize every tool at its disposal to reach Ukrainian capitulation. NewsFront and SouthFront have put forth AI-assisted

misinformation and continue to reach (potentially) sympathetic audiences to lessen political opposition to Russia. These sites have been singled out as two of the most disruptive, as shown through some of their headlines (e.g. “Ukrainian-American Biolabs’ Existence Proved by Various Sources” mentioned above) as well as their accounts that push for false narratives. The State Department decided to label them as disinformation outlets for these reasons, but such a process is not always clear-cut. Due to the opaque nature of Russian news, deciphering the true source and intent of mis/disinformation can prove difficult. While this is simply a contemporary research case study in 2022, these sites provoke a plethora of questions and considerations for what is becoming known as “the war for information.”

The war for information and conclusion

A headline by the Canadian Global Times sums up much of the fear for future communication: “We are not prepared[.]”²⁸ While (political) propaganda has existed for thousands of years, it takes a gargantuan leap forward when paired with AI technology. Despite formidable obstacles for fabricating programmers and parties, the relative success of Russian endeavors at home and in Ukraine speaks volumes. Hurdles for seeking truth rise, as the individual Internet consumer must equip him- or herself with the desire to take the time and clarify sketchy or conflicting news. “Generative media” outpaces manual efforts to stop it and continues to grow in sophistication with technical advances and ever-increasing data training set sizes.²⁹ Is our informational world then destined for constant uncertainty and meticulous fact-checking? The prevalence of “fake news” seems to answer in the affirmative. However, modern consumers may consider themselves lucky, as they live in an era in which fact-checking is at least possible, if not widely available. Previous periods afforded little of such luxury to anyone. The heart of the problem then lies in the relationship between these high-powered technologies and the individual.

The understanding that technology is central to our world and daily lives is not a profound one, but it has been commented on by several famous thinkers. José Ortega y Gasset and Hannah Arendt comment at length about the great potentialities for “technicism” in their respective works, *The Revolt of the Masses* and *The Threat of Conformism*. They explain the prevalence of technology as it is used to connect people and beliefs, both in these works and others. While neither could have predicted the explosion of AI onto the scene, both having written in the early- to late-20th century, they discussed instead the dangers of technology that produce “mass-men,” those who simply follow the information made available to them.³⁰ They both feared a decline in active reflection and citizen participation—that exponentially advancing technologies across sectors would diminish the demands people put upon themselves for such participation. Blissful ignorance loses its luster when one remembers that even inaction aids parties like Russia, who aim for minimal opposition to their cause.

These philosophers and President Zelenskyy alike call for action—the former for intellectual and citizen-related efforts, the latter for financial and military aid. Supporting Ukraine’s cry for help now, Ortega y Gasset and Arendt’s distinction should also be emphasized in the long run. A reflective, motivated, and active global citizen is better prepared to combat disinformation with all its hazardous implications. AI informational manipulation of individuals would certainly be a rigorous test for both philosophers’ opinions, yet it must be our hope that the 21st-century consumer is up to the challenge.

Endnotes

1. David Leonhardt, "[An Unlikely Hero](#)," New York Times, 14 March 2022.
2. Mike Whitney, "[The Man Who Sold Ukraine](#)," SouthFront.
3. "Press Freedom Index 2011–2012". Reporters Without Borders. Archived from the original on 31 January 2012.
4. NewsFront Website [home page](#).
5. Bobby Allyn, "[Researchers: Nearly Half Of Accounts Tweeting About Coronavirus Are Likely Bots](#)," NPR, 20 March 2020.
6. Dave Johnson, "[What is a deep fake? Everything you need to know about the AI-powered fake media](#)," Insider, 22 January 2021.
7. Tim Marcin, "[13 of our favorite deepfakes that'll seriously mess with your brain](#)," Mashable, 2 August 2020.
8. Georgios Petropoulos, "[The dark side of artificial intelligence: manipulation of human behavior](#)," Brügel, 2 February 2022.
9. Allyn, NPR.
10. "[Russia restricts access to several Western media websites](#)," Al Jazeera, 4 March 2022.
11. Carla Bleiker, "[Russia: Fewer independent media outlets every day](#)," Deutsche Welle, 3 March 2022.
12. Malaka Gharib, "[Not every war gets the same coverage as Russia's invasion — and that has consequences](#)," NPR, 4 March 2022.
13. Kathleen Doheny, "[Uncomfortably Numb: With COVID & Ukraine, Crisis Fatigue Thrives](#)," WebMD, 15 March 2022.
14. Kevin McSpadden, "[You Now Have a Shorter Attention Span Than a Goldfish](#)," Time, 14 May 2015.
15. John Villasenor, "[How to deal with AI-enabled disinformation](#)," The Brookings Institution, 23 November 2020.
16. Media Relations, "[New tool uses AI to flag fake news for media fact-checkers](#)," University of Waterloo, 16 December 2019.
17. "[Russia Profile - Media](#)," BBC, 8 June 2021.
18. Elizabeth Schulze, "[Russia just brought in a law to try to disconnect its internet from the rest of the world](#)," CNBC, 1 November 2019.
19. Kate Holton and Guy Faulconbridge, "[Global news media on defensive after Putin signs 'fake news' law](#)," Reuters, 4 March 2022.
20. Ibid.
21. "[Real news site or Russian propaganda?](#)" U.S. Embassy in Georgia.
22. U.S. Department of State, "[GEC Special Report: August 2020 Pillars of Russia's Disinformation and Propaganda Ecosystem](#)," August 2020.
23. U.S. Department of the Treasury press release, "[Treasury Sanctions Russians Bankrolling Putin and Russia-Backed Influence Actors](#)," 3 March 2022.
24. Aparna Banerjee, "[Digital war: How Russia is using deep fakes in Ukraine for propaganda](#)," Business Today India, 2 March 2022.
25. Ben Collins [Twitter thread](#), 28 February 2022.
26. Ben Collins and Jo Ling Kent, "[Facebook, Twitter remove disinformation accounts targeting Ukrainians](#)," NBC News, 28 February 2022.
27. U.S. State Dept.

28. Noor Ibrahim, “[‘We are not prepared’: Russia uses artificial intelligence, deep fakes in propaganda warfare](#),” Global News, 30 March 2022.
29. Renée DiResta, “[The Supply of Disinformation Will Soon Be Infinite](#),” The Atlantic, 20 September 2020.
30. José Ortega y Gasset, *The Revolt of the Masses*, Notre Dame, Indiana: University of Notre Dame Press, 1985.