



Volume 13 Issue 3



Blocking the information war? Testing the effectiveness of the EU's censorship of Russian state propaganda among the fringe communities of Western Europe

Christiern Santos Okholm European University Institute

Amir Ebrahimi Fard Maastricht University

Marijn ten Thij Maastricht University

DOI: <https://doi.org/10.14763/2024.3.1788>

Published: 29 July 2024

Received: 15 December 2023 **Accepted:** 23 May 2024

Competing Interests: The author has declared that no competing interests exist that have influenced the text.

Licence: This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License (Germany) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. <https://creativecommons.org/licenses/by/3.0/de/deed.en>
Copyright remains with the author(s).

Citation: Santos Okholm, C., Ebrahimi Fard, A., & ten Thij, M. (2024). Blocking the information war? Testing the effectiveness of the EU's censorship of Russian state propaganda among the fringe communities of Western Europe. *Internet Policy Review*, 13(3). <https://doi.org/10.14763/2024.3.1788>

Keywords: Censorship, Propaganda, Information war, Platforms, Russia

Abstract: In response to Russia's full-scale invasion of Ukraine in February 2022, the European Union banned or geo-blocked Russian propaganda media fearing the effects of the Kremlin's information warfare on internal public opinion during the largest security crisis in modern history. We investigate the blocks' effectiveness in limiting the sharing of Russian propaganda media content among vulnerable Western European fringe communities. By studying posting patterns on Facebook three months before and after the geo-block, we find that the geo-block successfully reduced the sharing of Russian propaganda media content among fringe communities and did not increase the sharing of other non-banned pro-Russian media. Furthermore, we found the geo-block increased sharing of content from alternative platforms, while the share of pro-Russian content doubled among these posts. These findings show the effectiveness of censorship in limiting foreign influence campaigns on major platforms, but they also show how alternative platforms allow for the continued spread of banned content.

Introduction

In the wake of Russia's full-scale invasion of Ukraine on February 24, 2022, the Council of the European Union (Council Regulation 2022/350) introduced a censorship regime against Russian propaganda media to protect the cohesion and stability of the Union and its member states during the largest security crisis in recent European history. Popularised as the geo-block, on March 2 the EU prohibited the broadcast, transmission, and distribution of content by Sputnik and Russia Today (henceforth RT) within the EU's borders, due to their role in pro-Russian disinformation campaigns (Ramsay & Robertshaw, 2019). Legitimised as a self-preserving countermeasure to Russia's doctrine of subverting adversarial liberal democracies (Galeotti, 2019; German, 2020; Kragh et al., 2020), the decision was an extraordinary act of censorship that clashed with liberal principles of press freedom. The ban of media content not only clashed with the EU fundamental rights charter paragraphs on freedom of expression but also the European Media Freedom Act's specifications on the protection of media content online, which were proposed seven months later.¹ Trapped between such core liberal values and the dangers of Russian subversion, it is prudent to ask if the EU's geo-block was effective in limiting Russian propaganda from reaching fringe communities?

Similar to previous work on democratic discontent (Rhodes-Purdy et al., 2023) and conspiracy theories (Imhoff & Bruder, 2014), we refer to fringe communities as communities that can broadly be thought of as communities who feel and perceive themselves as marginalised and disenfranchised within the broader society and who channel their frustration through radical ideologies (e.g. far-right/left) and conspiratorial beliefs to push back on what they perceive as a corrupt democratic system. As such, fringe communities can ascribe to various ideologies and beliefs, which from the perspective of influence campaigns remain relevant as long as their beliefs can be strategically exploited.

Russian elites regard influence campaigns as central in their asymmetrical confrontation with the West (Adamsky, 2017; Kovaleva, 2018), as they intend these influence campaigns to undermine the cohesion and resolve of strategic adversaries (Breitenbauch & Byrjalsen, 2019; Giles, 2019). With the large-scale invasion of Ukraine and as part of its hybrid warfare strategy, Russian dependency on undermining European resolve and support to Ukraine through subversive influence campaigns was only elevated. According to Walker et al. (2020) Russian influence

1. Charter of Fundamental Rights of the European Union, 389-405 (2016). European Media Freedom Act (Regulation 2024/1083).

campaigns are about exploiting existing societal grievances and tensions to undermine societal cohesion and influence political decisions. As a consequence, these frequently absorb themes from far-right (Vejvodová et al., 2017; Wagnsson & Barzanje, 2021), conspiracy movements (Dubow et al., 2021; Yablokov, 2015), Eurosceptic and anti-elitist (Lucas & Pomeranzev, 2016; Smoleňová et al., 2017) narratives, and amplify racial tensions (Freelon et al., 2020; Stewart et al., 2017) among Western audiences. By doing so, they exploit the frustrations and misperceptions among fringe communities.

Previous studies have found that extremist, strong ideological beliefs and acceptance of one conspiracy theory makes individuals more receptive to new false information (Ecker et al., 2022; Nemr & Gangware, 2019; Uscinski et al., 2020) and sharing of false information is done to empower the disenfranchised (Bang Petersen et al., 2018; Freelon et al., 2020; Young, 2021). The existing frustration, feelings of exclusion, and democratic discontent in fringe communities make them a primary audience for influence campaigns as these feelings are easily manipulated by foreign actors. By being more receptive and disproportionately targeted, fringe communities have become the main audience of Russian propaganda media and the geo-block can be seen as an attempt to stop Russian influence campaigns from reaching these communities by removing Sputnik and RT from the digital infrastructure accessible to European audiences. While an audit by Glazunova et al. (2023) found this implementation to be uneven, the geo-block's actual effectiveness in limiting the reach and sharing of Russian propaganda media, in particular among audiences most likely to be swayed by it, remains unknown/unanswered.

To investigate the effect of the geo-block's ability to stop fringe communities from consuming Russian propaganda media and disinformation, we turn to the academic literature on censorship. While censorship influences the cost-benefit analysis for audiences, it is unclear the degree to which it increases the demand of banned content or increases the access costs of pursuing that demand. Though the geo-block introduced considerable friction to consuming RT and Sputnik (Golovchenko, 2022; Pan & Roberts, 2020), it could also increase the sharing of other non-banned pro-Russian media (Pan & Siegel, 2019; Wong & Liang, 2021). This picture is further complicated by the number of alternative platforms that have emerged in recent years, allowing users to share and consume banned content across platforms (Horta Ribeiro et al., 2021; Mekacher et al., 2023; Rauchfleisch & Kaiser, 2021), thereby easing the imposed friction. Hence we ask how the geo-block affected the fringe communities' sharing of content originating from banned and non-banned Russian state media. Did it influence the sharing of content originat-

ing from alternative platforms and was this in-turn driven by pro-Russian content in regard to the full-scale invasion of Ukraine?

We investigate these expectations by mapping the online fringe communities of Western Europe and analysing their sharing of URLs associated with Russian propaganda media. Based on an extensive literature review of government, journalistic and academic reports of misinformation, and alternative media we create a list of fringe media and Russian propaganda media. We then use a snowball sampling method to identify public Facebook groups that regularly post and share URLs from our fringe media list. Next, we analyse these groups' posting patterns three months before and after the geo-block came into effect. Finally, we further determine the most commonly shared alternative platforms and subsequently study the content of these platforms within the same period.

Censoring propaganda

The scholarly debate on censorship provides us with conflicting answers to how audiences of unwanted information react to bans. Though most scholars subscribe to a cost-benefit approach, differences remain in how censorship affects the demand for banned information and the costs of pursuing these demands.

By being publicly announced, the geo-block could have had what Ong (2019) refers to as a *chilling effect* on the demand for pro-Russian content, as government actions socially stigmatise and threaten consumers of unwanted information (Li et al., 2018). However, as RT and Sputnik were stigmatised as Russian propaganda at the time of implementation, and the geo-block did not include risks to consumers, little indication supports the chilling of fringe groups. Indeed, Pan and Roberts (2020) found that frequent consumers would be the least affected by Chinese government censorship on Wikipedia and still go directly to banned pages, while incidental consumers who accessed pages via advertisement were most affected by it. However, as the ban did deteriorate the accessibility to RT and Sputnik's content considerably, i.e. prohibiting the sharing of the two's content on social media, it created what Roberts (2020) refers to as friction. As audiences can evade bans on digital media with Virtual Private Networks, the geo-block made it more difficult and strenuous for regular consumers to access RT and Sputnik. Studying a similar counter-disinformation measure, Golovchenko (2022) found that the Ukrainian government's censorship of Russian social media platforms did not stop but did heavily decrease user activity from Ukraine. As such, there are some indications of the geo-block being effective, providing us with our first hypothesis.

- H1: Sharing of content originating from RT and Sputnik fell post geo-block.

Contrary to this position, scholars have argued that censorship may increase sharing behaviour by highlighting issues in the public eye, i.e. the Streisand Effect (Jansen & Martin, 2015). Pan and Siegel (2019) found the Saudi government's imprisonment of dissidents backfired by drawing more online attention to dissidents' causes and boosting activism. According to Wong and Liang (2021), this can be explained by the inverted inferences that citizens make of censorship, such as interpreting censorship as proof of truth. In their survey experiment, government censorship eroded citizens' trust in the government and increased the credibility of banned information. This can be expected to be more pronounced among the fringe communities that Russian disinformation caters to, as they view alternative belief systems and narratives as ways to rebel against what they perceive as an unjust political system and counter-weight dominating narrative (Holt, 2018; Imhoff & Bruder, 2014). While RT and Sputnik were banned, several other Russian state media and disinformation outlets remained active and increasingly began parroting Russian state propaganda following the invasion, making pro-Russian disinformation accessible. Hence it can be expected that the ban increased the demand for pro-Russian media not included by the geo-block.

- H2: Sharing of content originating from pro-Russian media increased after the geo-block.

In extension of this, a growing literature argues that friction of accessibility in turn increases the demands for ways to circumvent imposed restrictions and reduce friction. As pressure to remove illegal and harmful content has increased on traditional platforms, e.g. Facebook and YouTube (who both enforced the geo-block), an ecology of alternative platforms emerged. Taking a free-speech absolutist position, these alternative platforms offer a safe haven for unwanted information (Nouri et al., 2021; Trujillo et al., 2020). A number of recent studies, e.g. Horta Ribeiro et al. (2021), Aliapoulios et al. (2021), Mekacher et al. (2023), and Rauchfleisch and Kaiser (2021) show that both the influencers and their audiences migrate to alternative platforms in response to platform censorship. As both RT and Sputnik announced setting up profiles on alternative platforms ahead of the geo-block, e.g. Gab, Odysee, Rumble, Telegram, and VKontakte (RT, 2022; Sputnik, 2022), it is not unlikely that sharing patterns were directed towards these platforms. Due to the growing plethora of alternative platforms (on which RT and Sputnik were active following the ban), platform migration does not only entail the successful transfer of user activity but also coordinating the direction of which platforms to move to. While the above-mentioned scholars focus on completed platform migration by

comparing general user activity on a platform or cross-referencing accounts, they fail to catch when communities become fragmented on multiple platforms and are vulnerable to the exact time when communities consolidate around a new platform when defining a completed platform. Instead, we suggest looking at cross-platform sharing between potential new platforms and established communities as an indicator of communities' attempts to meet the demand of circumventing censorship and dealing with this coordination problem. Specifically, the preferred platform to migrate to can be signalled by sharing links to specific fora on alternative platforms and content from specific actors' accounts or channels within well-established fora used by fringe communities. If censorship increases the demand for alternative platforms to alleviate the friction created by an intervention (e.g. geo-block), a reflection in levels of cross-platform sharing can be expected to precede a complete migration. Hence this gives us our third hypothesis.

- H3: Levels of cross-platform sharing within the community increase post geo-block.

However, while cross-platform sharing may indicate migration, this may be driven for other reasons. Systemic censorship of anti-vaxx and far-right content persists and may be more relevant than pro-Russian content for fringe communities. Hence an increase in cross-platform sharing may also express a general trend of platform migration which is not a result of the geo-block of Russian propaganda media. To verify cross-platform sharing is driven by the geo-block, we investigate if the shared content is related to the geo-block, i.e. being pro-Russian. As the reason for the geo-block was to halt the diffusion of pro-Russian frames, which may undermine European public opinion and decision-making regarding Russia's invasion of Ukraine, a Streisand Effect, or what Wong and Liang (2021) refer to as inverted inferences, may lead to increased cross-platform sharing of pro-Russian content. Hence, the geo-block will lead to popular alternative platforms having more pro-Russian content. We therefore present a sub-hypothesis.

- H3.1 : Popular alternative platforms will see an increase in pro-Russian content.

Method

To explore the geo-block's effect on the online reach of Russian propaganda media and content among fringe communities in Europe, we rely on the previously gathered data set by Santos Okholm et al. (2024) , as it allows us to compare posts within this community and focus on audiences who consume banned content regularly. The dataset maps fringe communities' presence on Facebook's public groups,

done through an automated snowball sampling starting from a list of 202 Western European fringe media. We use the term fringe media mainly with methodological intentions as a label for media that continuously share content that can be labelled as misinformation, conspiracy theories, and radical beliefs. It draws inspiration from the argument by Holt (2018) that certain alternative media perform an important role for the wider fringe community as these alternative media empower fringe communities to push back on a perceived corrupt and dysfunctional system through spreading misinformation. This seeding list of fringe media was gathered through an extensive literature review of government, journalistic, and academic reports in the fall of 2021 and subsequently verified by national fact-checkers.² In April 2022, we used the Facebook tool CrowdTangle to collect all public Facebook groups that posted or shared links to any of the listed fringe media more than once a week throughout 2021. This resulted in a list of 492 public Facebook groups who regularly share fringe media (i.e. fringe groups). We then manually verified whether these groups both appeared organic and corresponded to the country of origin, e.g. by analysing content, posting patterns and group administrators' location. As most reports and fact-checkers were focused on media outlets that were either sharing far-right, anti-vaccine, or conspiracy theory content, the identified groups largely fell into such categories, with only a few being far-left. Countries included are: Austria (96), Germany (139), France (77), Spain (78), Italy (37), Netherlands (14), Denmark (2), Sweden (36) and Finland (13). The data set brings with it methodological choices in need of clarification. Though only covering a fraction of the EU's member states, it provides useful insight into the key strategic European adversaries for Russia (e.g. France and Germany) and alleviates the overemphasis within the disinformation literature on US and Eastern European cases. Secondly, while fringe communities are active on other platforms, Facebook remains central for Russian influence campaigns' efforts (Institute for Strategic Dialogue, 2022) and remains key for European audiences (Statista, 2023) and fringe communities (The Soufan Center, 2021).

The data set's ability to isolate frequent users of banned content further complements the literature on censorship, which often tends to infer audiences' reaction to bans by studying a general population's behaviour to bans of content they may not be aware of (Hobbs & Roberts, 2018; Stoycheff et al., 2018) or study the traffic to specific websites (Nabi, 2014; Pan & Roberts, 2020) without qualifying the origins of that access. This approach risks understating effects on most affected audiences and overstating behaviour by general audiences. This is further problematic

2. Fact-checkers were contacted via the International Fact Checking Network Poynter.

due to the nature of online disinformation campaigns, relying on computational amplification and manipulation of engagement metrics (Broniatowski et al., 2018; Kriel & Pavliuc, 2019) and the targeting of fringe communities discussed above. Instead, we focus on a subset of online users' reaction to content they consume regularly being banned, similar to the study by Golovchenko (2022) on Ukrainian Vkontakte-users' response to the Ukrainian ban of the platform.

To measure the effect of the geo-block, we compare the posting and sharing frequency of URLs related to our first three hypotheses three months before and after the enforcement on 2 March 2022. As members of fringe communities recommend and share links with their peers, we use these recommendations as indicators of the communities' self-driven content sharing. However, as the geo-block removed Sputnik and RT on the network layer (Keremoğlu & Weidmann, 2020), the EU required associated URLs to be inaccessible in Europe. However, as noted by Glazunova et al. (2023), Facebook's implementation was irregular and only included notification of content being banned in Europe but did not prohibit users from posting affected URLs or accessing via VPNs. As such, H1 will also be a testament to the degree to which Facebook enforced the geo-block.

To test H1, we collected a list of URLs of the different Western European language versions of RT and Sputnik (i.e. Spanish, English, German, French and Italian) and identified mirror sites (RT's English language version was accessible on both the official site "<https://www.rt.com>" and the mirror site "<https://swentr.site>") in the Spring of 2023. This resulted in a list of 45 URLs to RT and Sputnik websites. As both RT and Sputnik were active on social media prior to the geo-block, we collected a second list of the accounts that were advertised on the two media's various language versions. This resulted in a list of 84 accounts from Facebook, Twitter, Telegram, Vkontakte, Odysee, Rumble, TikTok, RuTube, LiveJournal, Odnoklassniki, Instagram, and Koo.

To test H2's expectation to the sharing of content originating from other Russian state media, we conducted a second literature review of intelligence, research, and journalistic reports on online propaganda media attributed or with strong pro-Russian to the Russian state, which were either in English or other Western European languages.³ We then collected the URLs of the various language versions of their websites. This provided a list of 185 unique URLs.

3. For instance, investigations by Crossover (2022) and Le Monde (Pernet, 2022) found the French and English platform Donbas Insider, which consistently provides pro-Russian propaganda on the war in Ukraine, has close ties with the Russian military and RT.

To test H3's expectations of alternative platforms, we collected a list of top domain URLs from previously identified platforms, i.e. Rumble, Bitchute, Gettr, Parler, Telegram, and Substack with Western origins (Center for Countering Digital Hate, 2022; Stocking et al., 2022), as well as non-Western platforms such as the Chinese TikTok, Russian VKontakte, and RuTube, which have been known to be used to promote propaganda by their autocratic governments (Dietrich, 2023; Meduza, 2022; Ryan et al., 2020).

Based on the findings on H3 of Telegram being the most popular alternative platform shared in fringe communities, we collected the top 50 most cross-platform shared Telegram channels before and after the geo-block came into effect to study the content of the channels. After an initial data analysis showed that these all operated in either German, French, Italian, or Spanish, we then collected the top 20 channels before and after the geo-block for the four language areas in our data set.⁴ We then reviewed the content of the channels between 2-20 March 2022 to identify the prevalence of channels spreading pro-Russian content about Ukraine. Based on the European Digital Media Observatory (n.d.) database of Russian frames on the war in Ukraine that are debunked by fact-checkers, we define pro-Russian content as Telegram posts that framed Ukraine as a Nazi-regime, as a Western puppet-regime, accused Ukraine of committing war crimes committed by Russian forces, framed the outbreak of the war as Western provocation, or claimed that Ukraine has weapons of mass destruction, e.g. biological or nuclear weapons, intended for use against Russia. If channels posted content that fits any one of these topics in the two weeks after the geo-block, they were categorised as pro-Russian. As channels' popularity may fluctuate and Telegram also enforced the geo-block and banned access to Sputnik and RT, we identify channels appearing in the top both pre- and post-ban as "continued" (see figure 4 and 5). Channels that were inactive at the time of analysis (July 2023) were categorised as such. Both lists of outlets, media, fringe groups, and Telegram channels are available upon request.

4. For German this was done by merging Austrian and German Facebook groups.

Analysis

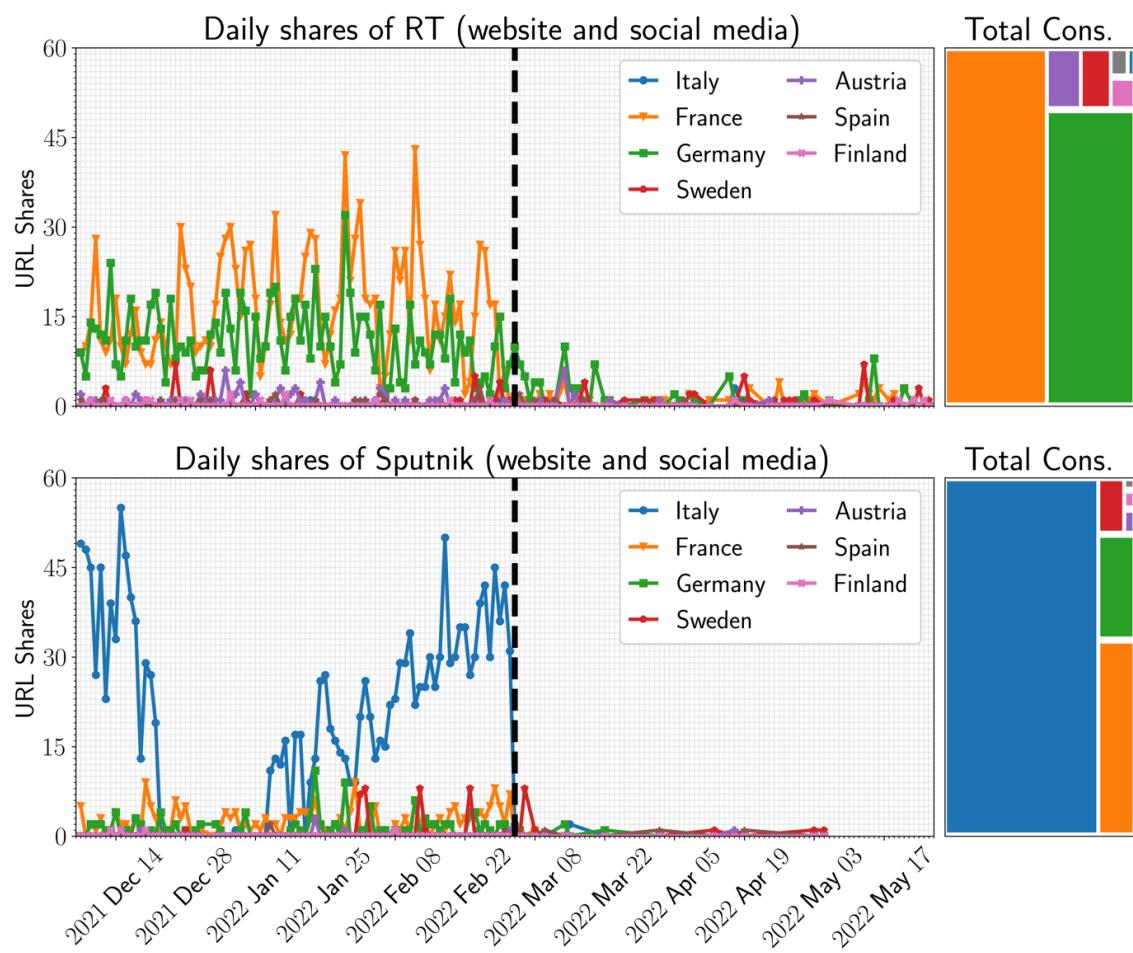


FIGURE 1: Historical analysis of daily sharing of URLs associated with RT (above) and Sputnik (below) among Western European fringe communities, three months before and after the enforcement of the EU's geo-block.

Based on the historical analysis in Figure 1, we confirm that sharing of content originating from RT and Sputnik fell sharply after the geo-block came into force (H1). Pre-geo-block, see that RT was popular among French and German fringe communities. While being the two largest samples in our data set, this can also be linked to RT being available in the two languages and having gained previous popularity among these national fringe communities.⁵ As RT does not have an Italian version but Sputnik does, we see a similar pattern among the Italian community for Sputnik, whose sharing activity drops over the Christmas holidays. Post-geo-block we see occasional sharing of RT and Sputnik content, indicating the possibility of continued sharing despite Facebook's enforcement of the ban. However, as

5. RT has been effective in building relations to German far-right voters (Berzina et al., 2021) and French yellow-vest protestors (AVAAZ, 2019).

the volumes are far lower than previous levels, the ban did increase the friction of accessing banned content and has proved effective in limiting the sharing of RT and Sputnik URLs among these communities, indicating also that Facebook did enforce the ban.

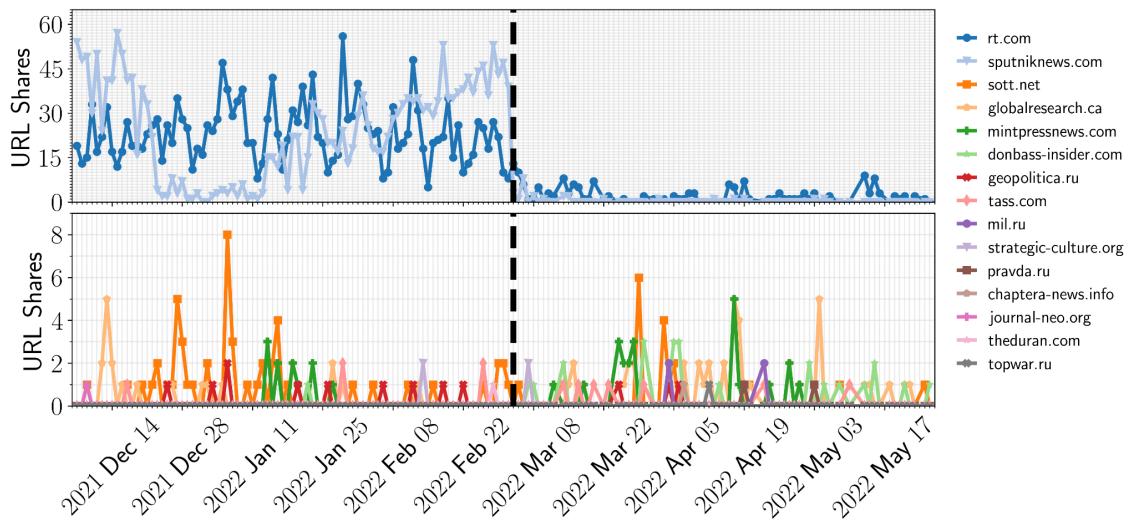


FIGURE 2: Online shares of Russian propaganda media in Western European fringe communities. Historical analysis of daily sharing of URLs associated with Pro-Russian media (see to the left) among Western European fringe communities, three months before and after the enforcement of the EU's geo-block.

In Figure 2, we see limited support for H2 as the sharing of content originating from pro-Russian media among Western European fringe communities on Facebook was largely unaffected by the geo-block. A reason for this lack of Streisand Effect may be due to the relative anonymity of the identified Pro-Russian media pre-geo-block, which is also reflected in the fact that post-block volumes mainly consist of RT URLs. Although some non-banned pro-Russian outlets, like Donbass Insider, Sign of the Times/Essence of the Times (Carrasco Rodrigues, 2020), Global Research (DiResta & Grossman, 2019) and the Mint Press News (Rudolph & Morley, 2020; Zawadzka, 2018) saw an increase of content sharing after the ban, they remained peripheral within the broader fringe community.

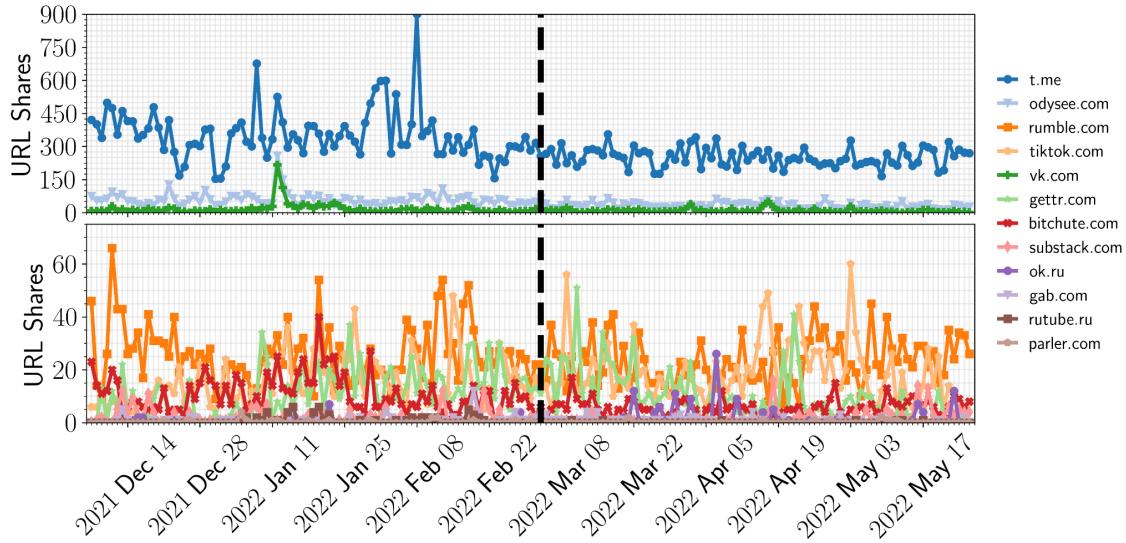


FIGURE 3: Shares of alternative platforms in Western European fringe communities. Historical analysis of daily sharing of URLs of alternative social media platforms (see to the left) among Western European fringe communities, three months before and after the enforcement of the EU's geo-block.

As shown in Figure 3, the geo-block of RT and Sputnik did not lead to an increase of cross-platform sharing content originating from alternative platforms, which remained high throughout the period. While the overwhelming majority of shared content was from private-owned Telegram, the popularity of private video-hosting sites like Odysee and Rumble and the Chinese state-controlled TikTok indicate a diversification of platforms within the fringe community. Although Figure 1 showed that RT and Sputnik were successful in catering to fringe communities, Figure 3 shows that they were not as integral members of these communities as to prompt an increase in sharing of alternative platforms and initiate a platform migration process.

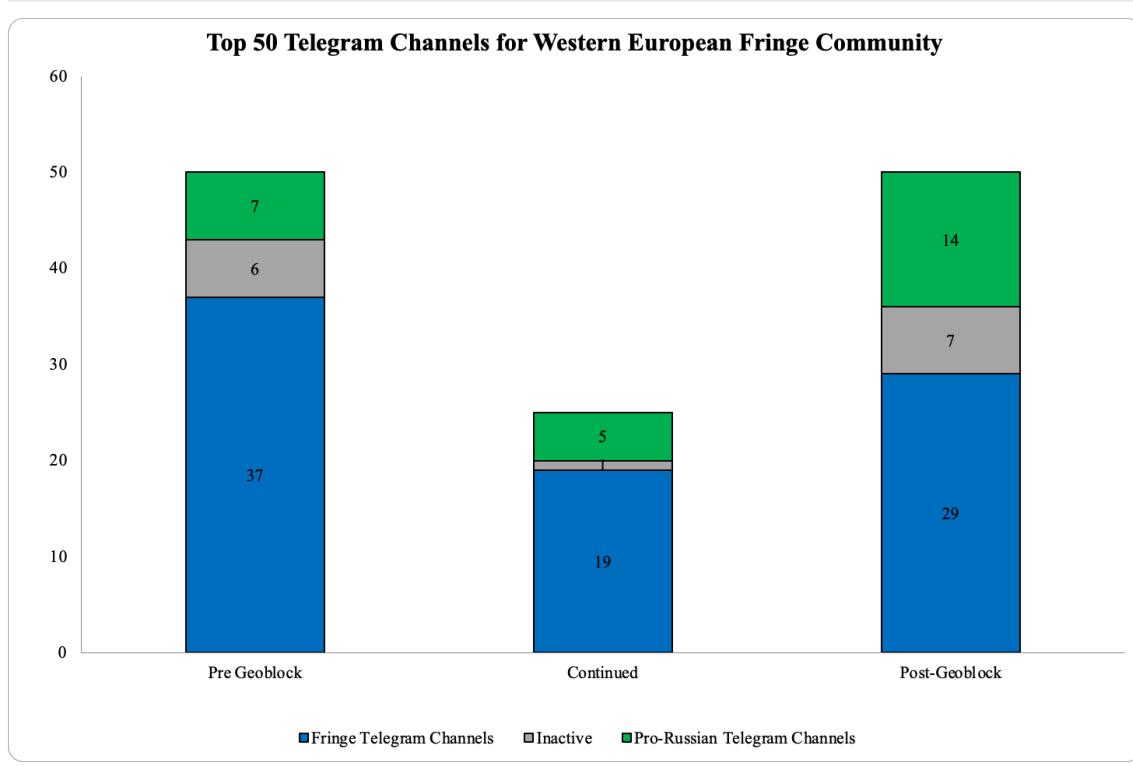


FIGURE 4: Top 50 Telegram channels pre- and post-geo-block among Western European fringe communities. Channels were categorised as inactive, fringe, or pro-Russian based on the prevalence of pro-Russian disinformation post-geo-block from March 2-20, 2022.

Based on the previous analysis, we explore the nature of the channels on the most popular alternative platform, Telegram, within the fringe community pre- and post-geo-block (see Figure 4). Although a majority of the analysed channels promoted far-right, anti-vaxx, and anti-establishment content, we find that pro-Russian Telegram channels among the top 50 channels for the entire period doubled from 7 to 14, supporting H3.1. While two pro-Russian channels fell in popularity, one of which was RT France, which was banned by Telegram, five pro-Russian channels continued being active in the top 50 and were joined with eight new pro-Russian channels. These channels would share and post pro-Russian content on Ukrainian Nazism, Ukrainian/Western development of bio-weapons and fake evidence on orchestrating war crimes, along with accusing the West of provoking the war. The majority of the top 50 channels were German (post-geo-block 32), French (post-geo-block 11), Spanish (post-geo-block 3) and Italian (post-geo-block 1), reflecting the national popularity of Telegram and a bias in our data set. To further investigate if the pattern above resonates across these national differences, we conducted a similar analysis of the top 20 channels in the three language spheres.

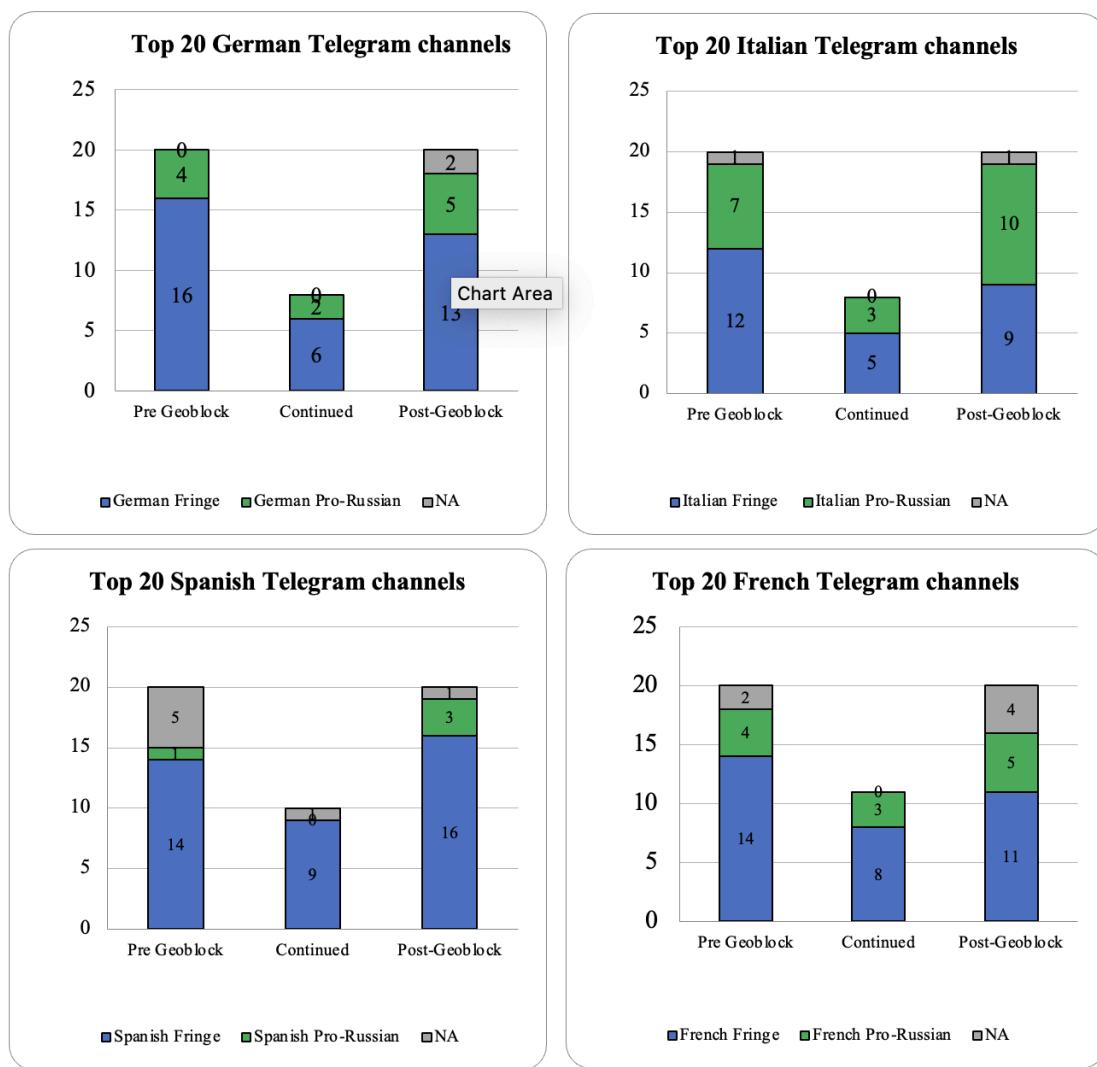


FIGURE 5: Overview of top 20 Telegram channels pre- and post-geo-block among German, French, Spanish and Italian fringe communities on Facebook. Channels were categorised as inactive, fringe or pro-Russian based on the prevalence of pro-Russian disinformation post-geo-block from March 2-20, 2022.

Looking into the national differences between the top 20 Telegram channels, we find that pro-Russian channels became more popular in German, Italian, French, and Spanish communities, although such channels were in a large minority. In line with the Streisand Effect, the most pronounced increase in popularity was found in Italy, this is moderated by the small number of times Telegram channels were shared in the Facebook groups. On average, Italian channels were shared 56.4 times post-geo-block among the Italian data set, while German (243 times) and French (109 times) were more popular. These, however, saw a more moderate increase from 4-5 pro-Russian channels, which would echo pro-Russian content on the Ukrainian Nazi State, war crimes, the Ukrainian usage/development of biological and nuclear weapons, and blame the invasion on Western provocation. This un-

derscores that the geo-block did not have a particular effect in stopping pro-Russian content from reaching Western European fringe communities, which would continue the sharing on Facebook via alternative platforms. However, it is important to note that pro-Russian Telegram channels would, on average, be shared fewer times than for the entire data set (29.3 for the Italian channels, 137 for the French, 187 for the German and 34.5 for the Spanish), moderating the Streisand Effect.

Conclusion

In sum, we find that the geo-block imposed considerable friction on the continued sharing of pro-Russian propaganda among fringe communities on Facebook, despite irregular implementation by Facebook. It effectively collapsed the extensive sharing of RT and Sputnik in Germany, France, and Italy without their position being superseded by other non-banned pro-Russian media, which remained peripheral in the broader community. At first sight, our findings lend credence to the democratic self-defence argument and the effectiveness of state censorship against subversive propaganda by other states.

The Streisand Effect assumes that, by imposing a ban on RT and Sputnik, one will draw more interest to these sources out of curiosity rather than the intended outcome of discouraging interest. As pre-ban primary audiences of unwanted content may already believe this to be true, bans could become promotional activities for other potential and unaware audiences who become aware of unwanted content after the implementation of the ban. Therefore, our results nuance this Streisand Effect as we show that this effect is only present for audiences that were unaware of the banned sources and not for the pre-ban primary audience.

However, since pro-Russian content became more popular on alternative platforms, there are indications of such platforms easing the friction imposed by bans and increasing the likelihood of a Streisand Effect even among the existing audiences. Although this effect was modest compared to the drop in sharing of propaganda media, it highlights the possibility of propaganda actors circumventing bans and may indicate that bans may at some level simply push content over to alternative platforms less likely to enact censorship. As alternative platforms are growing in number and becoming more popular among fringe communities, this poses an increased challenge to state censorship in the digital age. While the abundance of alternative platforms is likely to also create coordination problems for the wider community and thereby impose short-term friction, it is doubtful that censorship will maintain a long-term effect. The existence of such alternative platforms are

therefore likely to change the level of friction that censorship imposes and as a result affects audiences' cost-benefit analysis towards consuming unwanted information.

The findings present a nuanced picture of the EU's decision to geo-block Russian propaganda media and influence campaigns. While the ban was effectively implemented on a major digital platform and reduced the sharing of main propaganda platforms, alternative platforms did allow for some circumvention, questioning the degree to which the ban only was a substantial, yet temporary, setback for Russian influence campaigns. This finds some support in recent bans of an additional four Russian propaganda media implemented by the European Council (2024). Meanwhile, the increase of pro-Russian content and disinformation on Telegram underscores how dependent such bans are on private actors implementing them effectively. While Telegram indeed blocked Sputnik and RT channels from European users, similar content remained on the platform. The degree to which bans do increase the user base of alternative platforms is unclear, but the unwillingness of such platforms to implement digital censorship poses a challenge for such EU policies.

Moving forward, we stress that an act of censorship should not be taken lightly, as costs have to be weighed against core liberal values. Although our results do not show a large spill-over effect to alternative platforms, we do stress that we have only focused on the content on other platforms and did not include the magnitude of the spill-over of users. In our opinion, one cannot argue the successfulness of a ban of a particular source without a good understanding of the spill-over effects, both in terms of content and volume. We therefore call for future research to further explore these effects in more detail.

References

- Adamsky, D. (Dima). (2017). From Moscow with coercion: Russian deterrence theory and strategic culture. *Journal of Strategic Studies*, 41(1–2), 33–60. <https://doi.org/10.1080/01402390.2017.1347872>
- Aliapoulios, M., Bevensee, E., Blackburn, J., Bradlyn, B., De Cristofaro, E., Stringhini, G., & Zannettou, S. (2021). A large open dataset from the Parler social network. *Proceedings of the International AAAI Conference on Web and Social Media*, 15, 943–951. <https://doi.org/10.1609/icwsm.v15i1.18117>
- Avaaz. (2019). *Yellow Vests flooded by fake news. Over 100m views of disinformation on Facebook* (Report Version 1.2). <https://avaazimages.avaaz.org/Report%20Yellow%20Vests%20FINAL.pdf>
- Berzina, K., Blutguth, C., & Metzger, D. (2021). *Between messaging and manipulation: How Russia,*

China, Turkey, and Iran engaged in German political discourses in June 2021 [Report]. Alliance For Securing Democracy. <https://securingdemocracy.gmfus.org/between-messaging-and-manipulation-how-russia-china-turkey-and-iran-engaged-in-german-political-discourses-in-june-2021/>

Breitenbauch, H. Ø., & Byrjalsen, N. (2019). Subversion, statecraft and liberal democracy. *Survival*, 61(4), 31–41. <https://doi.org/10.1080/00396338.2019.1637118>

Broniatowski, D. A., Jamison, A. M., Qi, S., AlKulaib, L., Chen, T., Benton, A., Quinn, S. C., & Dredze, M. (2018). Weaponized health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate. *American Journal of Public Health*, 108(10), 1378–1384. <https://doi.org/10.2105/AJPH.2018.304567>

Carrasco Rodríguez, B. (2020). *Information laundering in Germany [Report]*. NATO Strategic Communications Centre of Excellence. <https://stratcomcoe.org/publications/information-laundering-in-germany/23>

Center for Countering Digital Hate. (2022, January 27). *Substack generates at least \$2.5 million in revenue from anti-vaccine newsletters per year [News]*. <https://counterhate.com/blog/substack-generates-at-least-2-5-million-in-revenue-from-anti-vaccine-newsletters-per-year/>

Charter of Fundamental Rights of the European Union. (2016). *Official Journal*, C202, 7 June, 389–405. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:12016P/TXT&rid=3>

Council Regulation 2022/350. (2022). *Council Regulation (EU) 2022/350 of 1 March 2022 amending Regulation (EU) No 833/2014 concerning restrictive measures in view of Russia's actions destabilising the situation in Ukraine*. Council of the European Union. <http://data.europa.eu/eli/reg/2022/350/oj>

Crossover. (2022). *Disinformation on Donbas is only a Google autocomplete away—An investigation to demonstrate how French speaking Belgians were hinted at searching for dubious sources when looking up the word ‘Donbass’ in the Google search bar [Investigation report]*. <https://crossover.social/disinformation-on-donbas-is-only-a-google-autocomplete-away/>

Dietrich, P. (2023). *The key player in Russia’s cybersphere* (Policy Brief 4; DGAP Analysis). German Council on Foreign Relations. https://dgap.org/system/files/article_pdfs/DGAP%20Analysis%20November.%20September_20_2023_20pp.pdf

DiResta, R., & Grossman, S. (2019). *Potemkin pages & personas: Assessing GRU online operations, 2014–2019* [White paper]. Freeman Spogli Institute for International Studies. <https://doi.org/10.25740/cv483mb5313>

Dubow, B., Lucas, E., & Morris, J. (2021). *Jabbed in the back: Mapping Russian and Chinese information operations during the COVID-19 pandemic [Report]*. Center for European Policy Analysis. <https://cepia.org/comprehensive-reports/jabbed-in-the-back-mapping-russian-and-chinese-information-operations-during-the-covid-19-pandemic/>

Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E. K., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, 1, 13–29. <https://doi.org/10.1038/s44159-021-00006-y>

European Council. (2024). *Russia’s war of aggression against Ukraine: Council bans broadcasting activities in the European Union of four more Russia-associated media outlets* [Press release]. <https://www.consilium.europa.eu/en/press/press-releases/2024/05/17/russia-s-war-of-aggression-against-ukraine-council-bans-broadcasting-activities-in-the-european-union-of-four-more-russia-associated-media-outlets/>

European Digital Media Observatory. (n.d.). *War in Ukraine: The fact-checked disinformation detected*

in the EU [Database]. <https://edmo.eu/war-in-ukraine-the-fact-checked-disinformation-detected-in-the-eu/>

Freelon, D., Bossetta, M., Wells, C., Lukito, J., Xia, Y., & Adams, K. (2020). Black trolls matter: Racial and ideological asymmetries in social media disinformation. *Social Science Computer Review*, 40(3), 560–578. <https://doi.org/10.1177/0894439320914853>

Galeotti, M. (2019). *Russian political war: Moving beyond the hybrid*. Routledge.

German, T. (2020). Harnessing protest potential: Russian strategic culture and the colored revolutions. *Contemporary Security Policy*, 41(4), 541–563. <https://doi.org/10.1080/13523260.2020.1757251>

Giles, K. (2019). *Moscow rules: What drives Russia to confront the West*. Brookings Institution Press. [ht tp://www.jstor.org/stable/10.7864/j.ctt20d87f8](http://www.jstor.org/stable/10.7864/j.ctt20d87f8)

Glazunova, S., Ryzhova, A., Bruns, A., Montaña-Niño, S. X., Beseler, A., & Dehghan, E. (2023). A platform policy implementation audit of actions against Russia's state-controlled media. *Internet Policy Review*, 12(2). <https://doi.org/10.14763/2023.2.1711>

Golovchenko, Y. (2022). Fighting propaganda with censorship: A study of the Ukrainian ban on Russian social media. *The Journal of Politics*, 84(2), 639–654. <https://doi.org/10.1086/716949>

Hobbs, W. R., & Roberts, M. E. (2018). How sudden censorship can increase access to information. *American Political Science Review*, 112(3), 621–636. <https://doi.org/10.1017/S0003055418000084>

Holt, K. (2018). Alternative media and the notion of anti-systemness: Towards an analytical framework. *Media and Communication*, 6(4), 49–57. <https://doi.org/10.17645/mac.v6i4.1467>

Horta Ribeiro, M., Jhaver, S., Zannettou, S., Blackburn, J., Stringhini, G., De Cristofaro, E., & West, R. (2021). Do platform migrations compromise content moderation? Evidence from r/The_Donald and r/Incels. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1–24. <https://doi.org/10.1145/3476057>

Imhoff, R., & Bruder, M. (2014). Speaking (un-)truth to power: Conspiracy mentality as a generalised political attitude. *European Journal of Personality*, 28(1), 25–43. <https://doi.org/10.1002/per.1930>

Institute Strategic Dialogue. (2022). *The murky origin story of #IstandwithRussia: How influencer networks proliferating across social media platforms spread pro-Kremlin narratives and hashtags*. <http://www.isdglobal.org/isd-publications/the-murky-origin-story-of-istandwithrussia/>

Jansen, S. C., & Martin, B. (2015). The Streisand effect and censorship backfire. *International Journal of Communication*, 9, 656–671. <https://ijoc.org/index.php/ijoc/article/view/2498>

Keremoğlu, E., & Weidmann, N. B. (2020). How dictators control the internet: A review essay. *Comparative Political Studies*, 53(10–11), 1690–1703. <https://doi.org/10.1177/0010414020912278>

Kovaleva, N. (2018). Russian information space, Russian scholarship, and Kremlin controls. *Defence Strategic Communications*, 4, 133–172. <https://doi.org/10.30966/2018.RIGA.4.5>.

Kragh, M., Andermo, E., & Makashova, L. (2020). Conspiracy theories in Russian security thinking. *Journal of Strategic Studies*, 45(3), 334–368. <https://doi.org/10.1080/01402390.2020.1717954>

Kriel, C., & Pavliuc, A. (2019). Reverse engineering Russian Internet Research Agency tactics through network analysis. *Defence Strategic Communications*, 6(1), 190–227. <https://doi.org/10.3096>

6/2018.RIGA.6.6.

Li, Y., Yang, S., Chen, Y., & Yao, J. (2018). Effects of perceived online–offline integration and internet censorship on mobile government microblogging service continuance: A gratification perspective. *Government Information Quarterly*, 35(4), 588–598. <https://doi.org/10.1016/j.giq.2018.07.004>

Lucas, E., & Pomeranzev, P. (2016). *Winning the information war: Techniques and counter-strategies in Russian propaganda* [Report]. Center for European Policy Analysis. <https://www.lse.ac.uk/iga/assets/documents/arena/archives/winning-the-information-war-full-report-pdf.pdf>

Meduza. (2022, February 9). *Resurrecting Rutube: The Russian authorities have been investing in domestic ‘alternatives’ to YouTube, investigative journalists report*. <https://meduza.io/en/feature/2022/02/09/resurrecting-rutube>

Mekacher, A., Falkenberg, M., & Baronchelli, A. (2023). *The systemic impact of deplatforming on social media* (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2303.11147>

Nabi, Z. (2014). Resistance censorship is futile. *First Monday*, 19(11). <https://doi.org/10.5210/fm.v19i11.5525>

Nemr, C., & Gangware, W. (2019). *The complicated truth of countering disinformation* [Commentary]. War on the Rocks. <https://warontherocks.com/2019/09/the-complicated-truth-of-countering-disinformation/>

Nouri, L., Lorenzo-Dus, N., & Watkin, A.-L. (2021). Impacts of radical right groups' movements across social media platforms – A case study of changes to Britain First's visual strategy in its removal from Facebook to Gab. *Studies in Conflict & Terrorism*, 1–27. <https://doi.org/10.1080/1057610X.2020.1866737>

Ong, E. (2019). Online repression and self-censorship: Evidence from Southeast Asia. *Government and Opposition*, 56(1), 141–162. <https://doi.org/10.1017/gov.2019.18>

Pan, J., & Roberts, M. E. (2020). Censorship's effect on incidental exposure to information: Evidence from Wikipedia. *Sage Open*, 10(1). <https://doi.org/10.1177/2158244019894068>

Pan, J., & Siegel, A. A. (2020). How Saudi crackdowns fail to silence online dissent. *American Political Science Review*, 114(1), 109–125. <https://doi.org/10.1017/S0003055419000650>

Pernet, M. (2022, June 8). Les influenceurs prorusses en ordre de bataille [The pro-Russian influencers in battle mode]. *Le Monde*. https://www.lemonde.fr/pixels/article/2022/06/08/guerre-e-n-ukraine-les-influenceurs-prorusses-en-ordre-de-bataille_6129421_4408996.html

Ramsay, G., & Robertshaw, S. (2019). *Weaponising news: RT, Sputnik and targeted disinformation* [Report]. The Policy Institute, King's College London; <https://www.kcl.ac.uk/policy-institute/research-analysis/weaponising-news>

Rauchfleisch, A., & Kaiser, J. (2021). *Deplatforming the far-right: An analysis of YouTube and BitChute*. SSRN. <https://doi.org/10.2139/ssrn.3867818>

Regulation 2024/1083. (2022). *Regulation (EU) 2024/1083 of the European Parliament and of the Council of 11 April 2024 establishing a common framework for media services in the internal market and amending Directive 2010/13/EU (European Media Freedom Act)*. European Parliament and Council. <http://data.europa.eu/eli/reg/2024/1083/oj>

Roberts, M. E. (2020). Resilience to online censorship. *Annual Review of Political Science*, 23(1), 401–419. <https://doi.org/10.1146/annurev-polisci-050718-032837>

RT. (2022, March 4). *How to access RT.com*. <https://web.archive.org/web/20220304200814/https://www.rt.com/russia/551256-how-access-rt-censorship-bypass/>

Rudolph, J., & Morley, T. (2020). *Covert foreign money: Financial loopholes exploited by authoritarians to fund political interference in democracies* [Report]. Alliance For Securing Democracy. <https://securingdemocracy.gmfus.org/covert-foreign-money/>

Ryan, F., Fritz, A., & Impiombato, D. (2020). *TikTok and WeChat: Curating and controlling global information flows* (Policy Brief 37/2020). Australian Strategic Policy Institute. <https://www.aspi.org.au/report/tiktok-and-wechat>

Santos Okholm, C., Ebrahimi Fard, A., & ten Thij, M. (2024). Debunking and exposing misinformation among fringe communities: Testing source exposure and debunking anti-Ukrainian misinformation among German fringe communities. *Harvard Kennedy School Misinformation Review*, 5(1). <https://doi.org/10.37016/mr-2020-134>

Smoleňová, I., Chržová, B., Várenyiov, I., Fischer, D. a, Bartha, D. n, Deák, A. s, Rácz, A. s, & Turkowski, A. (2017). *United we stand, divided we fall: The Kremlin's leverage in the Visegrad countries*. Prague Security Studies Institute. https://www.ceid.hu/wp-content/uploads/2017/11/Publication_United-We-Stand-Divided-We-Fall.pdf

Sputnik. (2022). *Banned from reading Sputnik?* Sputnik News International. <https://sputnikglobe.com/20220819/banned-from-reading-sputnik-1099763143.html>

Starbird, K., Arif, A., Wilson, T., Van Koevering, K., Yefimova, K., & Scarneccia, D. (2018). Ecosystem or echo-system? Exploring content sharing across alternative media domains. *Proceedings of the International AAAI Conference on Web and Social Media*, 12(1). <https://doi.org/10.1609/icwsm.v12i1.15009>

Statista. (2023). *Number of users of selected social media platforms in Europe from 2017 to 2027* [Dataset]. <https://web.archive.org/web/20230509231225/https://www.statista.com/forecasts/1334334/social-media-users-europe-by-platform>

Stewart, L. G., Arif, A., Nied, A. C., Spiro, E. S., & Starbird, K. (2017). Drawing the lines of contention: Networked frame contests within #BlackLivesMatter discourse. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW), 1–23. <https://doi.org/10.1145/3134920>

Stocking, G., Mitchell, A., Matsa, K. E., Widjaya, R., Jurkowitz, M., Ghosh, S., Smith, A., Naseer, S., & St. Aubin, C. (2022). *The role of alternative social media in the news and information environment*. Pew Research Center. <https://www.pewresearch.org/journalism/2022/10/06/the-role-of-alternative-social-media-in-the-news-and-information-environment/>

Stoycheff, E., Burgess, G. S., & Martucci, M. C. (2020). Online censorship and digital surveillance: The relationship between suppression technologies and democratization across countries. *Information, Communication & Society*, 23(4), 474–490. <https://doi.org/10.1080/1369118X.2018.1518472>

The Soufan Center. (2021). *Quantifying the Q conspiracy: A data-driven approach to understanding the threat posed by QAnon* [Special report]. <https://thesoufancenter.org/research/quantifying-the-q-conspiracy-a-data-driven-approach-to-understanding-the-threat-posed-by-qanon/>

Trujillo, M., Buntain, C., & Horne, B. (2020, July 13). What is BitChute? Characterizing the “free speech” alternative to YouTube. *Proceedings of the 31st ACM Conference on Hypertext and Social Media*. HT '20, Orlando, FL. <https://doi.org/10.1145/3372923.3404833>

Uscinski, J. E., Enders, A. M., Klofstad, C., Seelig, M., Funchion, J., Everett, C., Wuchty, S., Premaratne,

K., & Murthi, M. (2020). Why do people believe COVID-19 conspiracy theories? *The Harvard Kennedy School Misinformation Review*, 1. <https://doi.org/10.37016/mr-2020-015>

Vejvodová, P., Janda, J., & Víchová, V. (2017). *The Russian connections of far-right and paramilitary organizations in the Czech Republic* [Country report]. Political Capital. https://politicalcapital.hu/pc-admin/source/documents/PC_NED_country_study_CZ_20170428.pdf

Wagnsson, C., & Barzante, C. (2021). A framework for analysing antagonistic narrative strategies: A Russian tale of Swedish decline. *Media, War & Conflict*, 14(2), 239–257. <https://doi.org/10.1177/1750635219884343>

Wong, S. H.-W., & Liang, J. (2021). Dubious until officially censored: Effects of online censorship exposure on viewers' attitudes in authoritarian regimes. *Journal of Information Technology & Politics*, 18(3), 310–323. <https://doi.org/10.1080/19331681.2021.1879343>

Yablokov, I. (2015). Conspiracy theories as a Russian public diplomacy tool: The case of Russia Today (RT). *Politics*, 35(3–4), 301–315. <https://doi.org/10.1111/1467-9256.12097>

Zawadzka, M. (2018). *Today's Potemkin village: Kremlin disinformation and propaganda in Poland* [Report]. Warsaw Institute. <https://warsawinstitute.org/todays-potemkin-village-kremlin-disinformation-propaganda-poland/>

Published by



in cooperation with

