CPH TECH POLICY BRIEF #1

SEPTEMBER 2022

PUBLISHED BY THE CROWN PRINCESS MARY CENTER, COPENHAGEN CENTER FOR SOCIAL DATA SCIENCE & DEPARTMENT OF POLITICAL SCIENCE / UNIVERSITY OF COPENHAGEN

INFORMATION CONTROL ON YOUTUBE AND RUSSIA'S INVASION OF UKRAINE

BY YEVGENIY GOLOVCHENKO

This first edition of the CPH Tech Policy Brief presents insights from the latest data memo by Yevgeniy Golovchenko, Kristina Aleksandrovna Pedersen, Jonas Skjold Raaschou Pedersen, and Anna Rogers. The research behind the data memo is supported by SODAS – Copenhagen Center for Social Data Science and the Department of Political Science, University of Copenhagen.

OVERVIEW

This CPH Tech Policy Brief summarises key insights from research on the informational aspect of Russia's invasion of Ukraine. The Brief dives deeper into the two attempts to control the flows of information on YouTube about the war:

- The recent Russian censorship law
- YouTube's own global ban on Russian state media channels as part of the platform's effort to improve the quality of information about the war

We show that YouTube's ban succeeded in reducing online engagement for the Russian state-controlled media, which serve as frequent sources of disinformation. However, our findings also suggest that some of that engagement may have moved to non-banned pro-Kremlin channels.

Overall, large tech firms are capable of curbing central disinformation sources when the policies are implemented globally. However, the broader media ecology needs to be taken much more into account, and policy dilemmas regarding online bans are bound to multiply in the future.

SOCIAL MEDIA AS PART OF THE WAR

We have seen a rising public interest in questions of "information warfare" tied to disinformation and propaganda. Lately, this has been fuelled by Russia's invasion of Ukraine.

Disinformation can be used to assist troops on the ground by sowing confusion and chaos[1]. Similarly, manipulative information can be used to mobilise or demobilise support for security policies and political parties[2,3,4,5].

Since Russia's annexation of Crimea in 2014, information warfare has been a central part of the Russian military approach[6]. However, information war does not only unfold through the production of information but also through information control[7].

Therefore, it is perhaps not surprising that social media platforms themselves have become important arenas where state and non-state actors fight for the "truth" about war events. This Brief focuses on YouTube as one such arena due to its status as one of the largest social media platforms in Russia, Ukraine as well as in the rest of the world.

INFORMATION CONTROL AS MEANS OF FIGHTING PROPAGANDA

In liberal democracies, content moderation, bans and censorship are increasingly used as common tools against propaganda and disinformation. The EU has for instance implemented a ban on the Russian state-controlled news outlets RT (former Russia Today) and Sputnik in response to their role in spreading disinformation about the Russian invasion of Ukraine in 2022 (The European Council).

The ban is unprecedented in scale in a European context as it is meant to be implemented in the entire EU. Following similar arguments, the UK also removed the broadcasting license for RT (The Independent). The Ukrainian authorities blocked access to Russian cable TV news channels briefly after Russia's first invasion of Ukraine in 2014 while also effectively implementing a large-scale ban on a list of Russian websites, including VKontakte, one of the most popular social media platforms in Ukraine[8]. In 2018, France passed a law against "fake news" that provided the legal grounds for removing false information and gave the French national broadcasting agency the power to suspend foreign state-controlled TV channels if they disseminated disinformation during three months before an election (Euronews). Social media platforms such as Facebook, Instagram, Twitter and YouTube followed suit by either removing false content, using algorithmic downranking, and labeling misleading information (The Guardian).

INFORMATION CONTROL ON YOUTUBE

Following Russia's invasion of Ukraine in 2022, Kremlin blocked access to Facebook, Twitter, and Instagram to limit alternative information about the war. On 4 March, the Russian government signed a law giving up to 15 years in prison for spreading "disinformation" about the "military special operation," i.e., information that contradicts the official narrative, including any reference to Russia's military actions as "war." Russian users are subject to the law even when leaving comments to content on social media that they have not uploaded themselves.

YouTube serves as a gateway of alternative news coverage from the world into Russia as one of the last non-blocked foreign social media platforms. However, Kremlin also uses YouTube to reach out to Russian-speaking audiences domestically and abroad (e.g., Ukraine, Latvia, Kazakhstan). On 11 March, as part of their anti-disinformation efforts, YouTube announced a global ban on Russian state media[9].

RESEARCH DESIGN AND DATA

In our latest data memo, we examined

- ➔ The effects of YouTube's own ban on online engagement
- ➔ The change in commenting engagement following Russia's censorship law
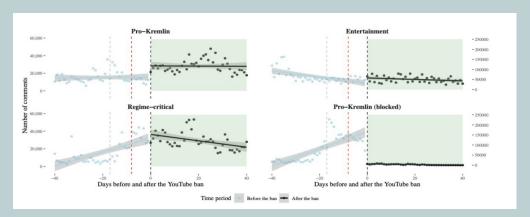
For this purpose, we collected ~12.3 million YouTube comments tied to nearly 14,000 videos from 40 Russian-speaking channels, including ten banned and ten non-banned pro-Kremlin channels, ten regime-critical channels, and 10 apolitical entertainment channels.

The exact timing of YouTube's own ban was not publicly announced well in advance. We can compare online engagement, measured as the number of comments, before and after YouTube's relatively sudden ban. Online engagement is essential for the channels because it helps boost the popularity of videos. As the Russian censorship law was announced well in advance of the date it took place, it is difficult to make a clear-cut causal interpretation of its effects on the online engagement on YouTube.

FINDINGS: YOUTUBE BAN AND THE RUSSIAN CENSORSHIP LAW

Our preliminary findings show that YouTube' ban reduced the number of comments for Russian-state media channels to almost zero just few hours after YouTube announced its policy (see bottom right on Figure 1). As a point of comparison, we see no decline in apolitical entertainment channels. While there is a decline in engagement with regime critical channels, that change is neither sudden (e.g., discontinuous) nor substantively large.

| FIGURE 1 | CHANGE IN NUMBER OF COMMENTS FOR BANNED PRO-KREMLIN MEDIA AND ENTERTAINMENT CHANNELS |

GREY LINE: INVASION (24 FEBRUARY) RED LINE: CENSORSHIP LAW (4 MARCH) BLACK LINE: YOUTUBE BAN (11 MARCH)

Source: Data memo by Yevgeniy Golovchenko, Kristina Aleksandrovna Pedersen, Jonas Skjold Raaschou Pedersen, and Anna Rogers.
FIGURE 1. Change in number of comments for banned Pro-Kremlin media and entertainment channels

However, we also observe a sharp increase in the number of comments for pro-Kremlin non-banned YouTube channels, as shown in the top left part of Figure 1. This is important because like the blocked media outlets, these channels also serve as sources of pro-Kremlin disinformation. This suggests that some users may simply have migrated from banned to non-banned pro-Kremlin channels.

Existing research suggests that uncertainty about what is allowed to be said and how is an effective censorship tool[10]. We expected that the Russian censorship law would prompt a significant decrease in comments on regime-critical channels due to its vagueness and severity because users would prefer to be on the safe side.

To our surprise, we see no decline in the number of comments on regime-critical channels after the censorship law was signed on 4 March (marked by the red dashed line in the bottom left part of Figure 1). We are currently examining whether the law may have changed the content of the comments rather than merely the number of comments. It is possible that opinions are now expressed through more careful language.

IMPLICATIONS

As this Brief illustrates, tech firms like YouTube are deeply entangled in modern conflicts. They have the technological means to make some news sources more visible than others – potentially influencing the flow of disinformation in the context of war. Indeed, our findings suggest that YouTube did successfully limit pro-Kremlin state media when they put in place a global ban in March 2022. There are, however, also challenges connected with such interventions.

The main takeaway is that we need to view the fight against disinformation in the context of a broader media ecology. Blocking one set of disinformation sources may not always be enough if users can easily switch to non-blocked sources that offer some of the same information as the blocked ones. Interventions that incentivise users to migrate across channels or even social media platforms may risk pushing users toward even more extreme disinformation sources.

This also raises another point: the need for more and better data. While scholars can evaluate attempts to control the flow of information by states or tech firms independently and relatively quickly, such studies are currently limited by a lack of access to data from YouTube and other social media firms offer even less access. Data access is necessary to inform societal debates about information control and modern conflicts. Without data, we risk discussing the issues without knowing how moderation, censorship, or online bans take place or what their intended or unintended consequences may be.

DILEMMAS

Online bans of disinformation raise a number of questions and dilemmas:

- To what extent and under which conditions should tech firms block disinformation sources?

- To what degree should transnational tech firms' interventions be regulated by states – and by which states? Is the answer different in the context of war?

- If one accepts the premise of banning disinformation sources, how broad should these bans be?

- And perhaps most importantly, how do we ensure transparency about the moderation process as well as the effects of such policies?

POLICY RECOMMENDATIONS

There are obvious potential efficiency gains to be had by using machine learning models, and based on this we have the following recommendations:

- Social media platforms need to view disinformation in the context of a broader media ecology. Blocking one set of disinformation sources may not work fully if users can easily switch to other non-blocked sources.

- Tech firm transparency and independent social data science research should go hand in hand to ensure a better understanding of the challenges that we will continue to face in a digitised society.

REFERENCES

1   THORNTON, R. (2015). THE CHANGING NATURE OF MODERN WARFARE: RESPONDING TO RUSSIAN INFORMATION WARFARE. THE RUSI JOURNAL 160(4), 40–48.

2   ZANNETTOU, S., T. CAULFIELD, E. DE CRISTOFARO, M. SIRIVIANOS, G. STRINGHINI, AND J. BLACKBURN. 2018. "DISINFORMATION WARFARE: UNDER-STANDING STATE-SPONSORED TROLLS ON TWITTER AND THEIR INFLUENCE ON THE WEB." ARXIV, JANUARY 28. HTTPS://ARXIV.ORG/ABS/1801.09288.

3   LINVILL, D. L., AND P. L. WARREN. 2019. "TROLL FACTORIES: MANUFACTUR-ING SPECIALIZED DISINFORMATION ON TWITTER." POLITICAL COMMUNICA-TION. PUBLISHED ELECTRONICALLY FEBRUARY . DOI:10.1080/10584609.2020 .1718257.

4   GOLOVCHENKO, Y., BUNTAIN, C., EADY, G., BROWN, M. A., & TUCKER, J. A. (2020). CROSS-PLATFORM STATE PROPAGANDA: RUSSIAN TROLLS ON TWITTER AND YOUTUBE DURING THE 2016 US PRESIDENTIAL ELECTION. THE INTERNATIONAL JOURNAL OF PRESS/POLITICS, 25(3), 357-389.

5   GOLOVCHENKO, Y., HARTMANN, M., & ADLER-NISSEN, R. (2018). STATE, MEDIA AND CIVIL SOCIETY IN THE INFORMATION WARFARE OVER UKRAINE: CITIZEN CURATORS OF DIGITAL DISINFORMATION. INTERNATIONAL AFFAIRS, 94(5), 975-994.

6   DARCZEWSKA, J. (2014). THE ANATOMY OF RUSSIAN INFORMATION WAR-FARE. THE CRIMEAN OPERATION, A CASE STUDY

7   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

8   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

9   REUTERS (2022). YOUTUBE ANNOUNCES AN IMMEDIATE BLOCK ON RUSSIAN STATE-FUNDED MEDIA CHANNELS GLOBALLY. EURONEWS.NEXT. ACCESSED MAY 21, 2022

10   STERN, RACHEL E., AND JONATHAN HASSID. 2012. "AMPLIFYING SILENCE: UNCERTAINTY AND CONTROL PARABLES IN CONTEMPORARY CHINA. COM-PARATIVE POLITICAL STUDIES 45 (10): 1230–54