Media and Communication
Open Access Journal | ISSN: 2183-2439
Volume 9, Issue 4 (2021)

Media Control Revisited: Challenges, Bottom-Up Resistance and Agency in the Digital Age

Editors
Olga Dovbysh and Esther Somfalvy

COGITATIO

Media and Communication, 2021, Volume 9, Issue 4
Media Control Revisited: Challenges, Bottom-Up Resistance and Agency in the Digital Age

Published by Cogitatio Press
Rua Fialho de Almeida 14, 2º Esq.,
1070-129 Lisbon
Portugal

Academic Editors
Olga Dovbysh (University of Helsinki, Finland)
Esther Somfalvy (Research Centre for East European Studies at the University of Bremen, Germany)

Available online at: www.cogitatiopress.com/mediaandcommunication

Table of Contents

COGITATIO

Media and Communication (ISSN: 2183–2439)
2021, Volume 9, Issue 4, Pages 1–4
https://doi.org/10.17645/mac.v9i4.4861

Editorial

Understanding Media Control in the Digital Age

Olga Dovbysh [1], Esther Somfalvy [2,*]

[1] Aleksanteri Institute, University of Helsinki, Finland; E-Mail: olga.dovbysh@helsinki.fi
[2] Research Centre for East European Studies at the University of Bremen, Germany; E-Mail: somfalvy@uni-bremen.de

* Corresponding author

Submitted: 1 September 2021 | Published: 21 October 2021

Abstract
Media control comprises multifaceted and amorphous phenomena, combining a variety of forms, tools, and practices. Today media control takes place in a sphere where national politics meet global technology, resulting in practices that bear features of both the (global) platforms and the affordances of national politics. At the intersection of these fields, we try to understand current practices of media control and the ways in which it may be resisted. This thematic issue is an endeavour to bring together conceptual, methodological, and empirical contributions to revise the scholarly discussion on media control. First, authors of this thematic issue re-assemble the notion of media control itself, as not being holistic and discrete (control vs freedom) but by considering it from a more critical perspective as having various modes and regimes. Second, this thematic issue brings a "micro" perspective into understanding and theorising media control. In comparison to structural and institutional perspectives on control, this perspective focuses on the agency of various actors (objects and subjects of media pressure) and their practices, motivations, and the resources with which they exert or resist control. Featuring cases from a broad range of countries with political systems ranging from democracy to electoral authoritarian regime, this issue also draws attention to the question of how media control relates to regime type.

Keywords
digital media; freedom of expression; internet governance; media pressure; media control; political regimes

Issue
This editorial is part of the issue "Media Control Revisited: Challenges, Bottom-Up Resistance and Agency in the Digital Age" edited by Olga Dovbysh (University of Helsinki, Finland) and Esther Somfalvy (Research Centre for East European Studies at the University of Bremen, Germany).

1. Introduction

Media control is discussed in relation to media and journalistic independence, freedom of information and expression worldwide. At the same time, media control comprises multifaceted and amorphous phenomena, made even more elusive as digital technologies blur the existing notions and create new ones about media control, its forms, and practices.

The elusiveness of the concept becomes visible in the terminology used to describe it: censorship, manipulation, instrumentalisation, influence, fraud, capture, pressure, discipline, or the "interference in journalistic autonomy" (Akhrarkhodjaeva, 2017; Goyanes & Rodríguez-Castro, 2019)—all of these terms are used by scholars and practitioners alike when they attempt to describe various aspects of control over mass media (Dovbysh & Mukhametov, 2020). However, one can hardly find a clear definition of what control over media is, and what it is not.

A full account of media control is not limited to pressure initiated by political forces but also considers economic (Herman & Chomsky, 1988/2010; Pleines, 2016) and social pressure, among others, resulting in different ways of agenda-setting, framing, and priming of media content. Moreover, the scholarly discussion of control by the state, "business," etc., should imply a nuanced study of different actors with their interests, roles, and positions in the media sphere.

Digitalisation has made the phenomenon even more multifaceted. Digital-born practices of media control like doxing or littering and the emergence of "trolls" or "buzzers" manipulate public discourse (for Turkey, see Baloğlu, 2021; for Russia, see Kiriya, 2021; for Indonesia, see Masduki, 2021). Journalists face "digital threats" or "digital violence" through surveillance, harassment, or data mining (Henrichsen et al., 2015). Together, they have led to a "mainstreamisation" of the digital space, meaning that state media outlets take over the digital space, which had previously mainly attracted oppositional discourses (for Russia, see Kiriya, 2021). The proliferation of artificial intelligence driven tools in the media sphere led to the emergence of new actors and tools of control in the form of technological corporations and the digital products and services they provide. The power of algorithms in the decisions that they make via prioritisation, classification, association, and filtering of data leads to a media dependency on algorithms and platforms and the leverage of platform power over journalistic practices and online information dissemination (Diakopoulos, 2019).

Simultaneously, new means of resistance and adaptation to various forms of control over and manipulation in the media have emerged. Media practitioners, no longer limited by newsrooms and institutionalised media outlets, have expanded their agency through the ability to produce and disseminate content via different channels and (social media) platforms in a hybrid and fragmented media sphere. Algorithmic power results in media practices of resistance (Velkova & Kaun, 2021) and adaptation to new forms of human-machine (inter)actions (on gatekeeping practices on the platforms, see Malinen, 2021).

The thematic issue puts forward an alternative approach to the scholarly discussion on media control in today's world in two ways. First, the authors re-assemble the notion of media control itself, as not being holistic and discrete (control vs freedom) but by considering it from a more critical perspective as having various modes and regimes.

As the idea of digital sovereignty has recently started to gain the attention of nation-states, internet governance has become a visible part of governments' activities to expand control over cyberspace. In her article, Anna Litvinenko (2021) explores Russia's strategic narrative on digital and internet sovereignty as a part of global internet governance. Based on document analysis and expert interviews, she reveals dependencies between narratives on internet policy and the elite's evaluation of the perceived benefits and threats of global connectivity. According to Litvinenko, the Russian case of internet sovereignty is an attempt to subject a highly decentralised network to tighten state regulation via a series of measures. Ilya Kiriya (2021) focuses on another means of state interference into the internet space—the isolation of oppositional discourses with the simultaneous creation of a massive flood of pro-state information. The author calls this strategy "littering the information space" (Kiriya, 2021, p. 23). Combining direct blocking measures and massive dissemination of state-funded online news, the strategy ensures state control and domination of pro-state discourses in the digital information space. Liudmila Sivetc and Mariëlle Wijermars (2021) present another mode of control in digital space in authoritarian regimes—the internet governance by its infrastructure—leveraging the power of private infrastructure owners to obtain control over content dissemination online.

Second, this thematic issue brings a "micro" perspective into understanding and theorising media control. In comparison to structural and institutional perspectives on media control, such as the ownership of media capital (Vendil Pallin, 2017) or the legal regulation of media, a "micro" perspective focuses on the agency of various actors (objects and subjects of media pressure) and their practices, motivations, and the resources with which they exert or resist control. What is the agency of journalists, media practitioners, or online activists under political pressure? What are the practices of resistance and strategies of adaptation? What are the actual challenges of media capture in the current technological environment?

Some contributions to this issue examine practices emerging in new media as technological advancement facilitates new forms of control. Uğur Baloğlu (2021) studies troll politics in Turkish Twitter, focusing on the ruling AK Party's politics on social media. In the example of the Boğaziçi University protests, the author examines how communication is suppressed through trolls and asks to what extent and how counter-trolls can intervene to create alternative discourses and shape public opinion. Masduki (2021) shows that in the case of Indonesia, the rise of digital media has resulted in new forms of control that target critical media outlets. It is characterised by the rise of non-state and societal control over digital media, where pressure is exerted by paid social-media buzzers who manipulate information and counter critical news regarding political leaders, contributing to journalist's self-censorship. Rashid Gabdulhakov (2021) explores the role of state-approved digital vigilantes in Russia. Based on the example of the *Hrushi Protiv* (Piggy Against) vigilante group, the author examines formations of citizens using digital media to expose "offences" carried out by fellow citizens. Such public shaming within online platforms allows the state to demonstrate a façade of civil society activism amid the silencing of certain types of critical publics while participants gain financial rewards and fame. Sanna Malinen (2021) studies volunteer moderation in Facebook groups. Practices of moderation not only shape public discussions in these groups but also regulate access to these discussions, which makes the moderators powerful though less visible gatekeepers of the digital public sphere.

Other authors focus on traditional media and examine how individual actors within such media resist pressure from their superiors or state actors. Marína Urbániková (2021) examines journalistic autonomy in the

Slovakian public service broadcaster and classifies resistance practices that journalists use to cope with perceived interference in their work by their media organisations. Esther Somfalvy and Heiko Pleines (2021) explore instances where journalists in Ukraine resisted censorship pressure during the Yanukovich presidency, asking which factors supported their agency. They find that while the nature of competitive authoritarianism does offer journalists opportunities for critical reporting, it is their individual characteristics—including professional ethics, networks, and job mobility—which defines whether and how the respective opportunities are used.

Finally, Jan Matti Dollbaum (2021) uses media under authoritarian conditions as data for research on protest events. Comparing protest event data from Russia that are based on different sources (English-language news agencies, dissident websites, local sources), the author demonstrates that while the data sources present different pictures of the protests, the divergence is systematic and can be put to productive use.

The cases featured in this volume come from a broad range of countries with political systems ranging from democracy to electoral authoritarian regime. Media outlets are among the first targets of governments that display authoritarian tendencies: The governments' attempts to maintain the image of a functioning democracy while tilting the political playing field in their favour leads to a variety of censorship practices (Levitsky & Way, 2010). This makes the media one of the major battlefields in political power struggles and draws attention to another dimension of media control that features prominently in the discussion, namely the question of how media control relates to regime type. Several studies have shown that in electoral or competitive authoritarian regimes, media manipulation is used more often than most other types of manipulation when regimes attempt to shift the "playing field" in their favour (Carter & Carter, 2018; Yeşil, 2018). State-funded digital violence as means of media control leads to the restriction of press freedom, which is linked to democratic backsliding (Freedom House, 2020a, 2020b). Populist and authoritarian governments instrumentalise communication within social media platforms, where polarising and "otherising" discourses are easy to create (Grinberg et al., 2019; Poell & van Dijck, 2014). At the same time, the platform companies themselves are fundamentally political actors that make political decisions. The algorithmic control exercised by the platform companies shapes the very notion of freedom of expression. The platforms' political influence, together with the low democratic accountability of their algorithms, lead to new facets and challenges of media control across political regimes. In sum, media control today occurs at the intersection of national politics with global technology, while heterogeneous practices of media control and the means of resisting it that we observe have emerged as a result of both features of the (global) technologies and the affordances of national politics.

Acknowledgments

This editorial was produced as part of the research project "Media Control as Source of Political Power: The Role of Oligarchs in Electoral Authoritarian Regimes," conducted by the Research Centre for East European Studies at the University of Bremen and received financial support from the Deutsche Forschungsgemeinschaft (DFG), grant No. 391270526. We are thankful to all participants of the Workshop on "Media Control as Source of Political Power in Central and Eastern Europe" in September 2019 at the Aleksanteri Institute, University of Helsinki, for their helpful feedback.

Conflict of Interests

The authors declare no conflict of interest.

References

Akhrarkhodjaeva, N. (2017). *Instrumentalisation of mass media in electoral authoritarian regimes: Evidence from Russia's presidential election campaigns of 2000 and 2008*. Ibidem.

Baloğlu, U. (2021). Trolls, pressure, and agenda: The discursive fight on Twitter in Turkey. *Media and Communication*, *9*(4), 39–51.

Carter, E. B., & Carter, B. L. (2018). Propaganda and electoral constraints in autocracies. *Comparative Politics Newsletter*, *28*(2), 11–18.

Diakopoulos, N. (2019). *Automating the news*. Harvard University Press.

Dollbaum, J. M. (2021). Protest event analysis under conditions of limited press freedom: Comparing data sources. *Media and Communication*, *9*(4), 104–115.

Dovbysh, O., & Mukhametov, O. (2020). State information contracts: The economic leverage of regional media control in Russia. *Demokratizatsiya*, *28*(3), 367–391.

Freedom House. (2020a). *Freedom on the net 2020: Indonesia*. https://freedomhouse.org/country/indonesia/freedom-net/2020

Freedom House. (2020b). *Freedom in the world 2020: Indonesia*. https://freedomhouse.org/country/indonesia/freedom-world/2020

Gabdulhakov, R. (2021). Media control and citizen-critical publics in Russia: Are some "pigs" more equal than others? *Media and Communication*, *9*(4), 62–72.

Goyanes, M., & Rodríguez-Castro, M. (2019). Commercial pressures in Spanish newsrooms: Between love, struggle and resistance. *Journalism Studies*, *20*(8), 1088–1109.

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 US presidential election. *Science*, *363*(6425), 374–378.

Henrichsen, J. R., Betz, M., & Lisosky, J. M. (2015). *Building digital safety for journalism: A survey of selected*

issues. UNESCO.
Herman, E. S., & Chomsky, N. (2010). *Manufacturing consent: The political economy of the mass media*. Random House. (Original work published 1988)
Kiriya, I. (2021). From "troll factories" to "littering the information space": Control strategies over the Russian internet. *Media and Communication*, *9*(4), 16–26.
Levitsky, S., & Way, L. A. (2010). *Competitive authoritarianism: Hybrid regimes after the Cold War*. Cambridge University Press.
Litvinenko, A. (2021). Re-defining borders online: Russia's strategic narrative on internet sovereignty. *Media and Communication*, *9*(4), 5–15.
Malinen, S. (2021). Boundary control as gatekeeping in Facebook groups. *Media and Communication*, *9*(4), 73–81.
Masduki. (2021). Media control in the digital politics of Indonesia. *Media and Communication*, *9*(4), 52–61.
Pleines, H. (2016). Oligarchs and politics in Ukraine. *Demokratizatsiya*, *24*(1), 105–127.
Poell, T., & van Dijck, J. (2014). Social media and journalistic independence. In. J. Bennett & N. Strange (Eds.), *Media independence: Working with freedom or working for free?* (pp. 182–201). Routledge.
Sivetc, L., & Wijermars, M. (2021). The vulnerabilities of trusted notifier-models in Russia: The case of Netoscope. *Media and Communication*, *9*(4), 27–38.
Somfalvy, E., & Pleines, H. (2021). The agency of journalists in competitive authoritarian regimes: The case of Ukraine during Yanukovich's presidency. *Media and Communication*, *9*(4), 82–92.
Urbániková, M. (2021). Resisting perceived interference in journalistic autonomy: A case study of public service media in Slovakia. *Media and Communication*, *9*(4), 93–103.
Velkova, J., & Kaun, A. (2021). Algorithmic resistance: Media practices and the politics of repair. *Information, Communication & Society*, *24*(4), 523–540.
Vendil Pallin, C. (2017). Internet control through ownership: The case of Russia. *Post-Soviet Affairs*, *33*(1), 16–33.
Yeşil, B. (2018). Authoritarian turn or continuity? Governance of media through capture and discipline in the AKP Era. *South European Society and Politics*, *23*(2), 239–257.

About the Authors

Olga Dovbysh is a postdoctoral researcher at the Aleksanteri Institute, University of Helsinki, and coordinator of the Russian Media Lab Network initiative. She works at the intersection of media studies, economic sociology, and political economy. Since January 2020, she has been working in the project Sustainable Journalism for the Algorithmic Future, which studies the challenges of algorithmic journalism in Russia and beyond.

Esther Somfalvy is a research fellow at the Research Centre for East European Studies at the University of Bremen. Her research interests include comparative authoritarianism studies, political institutions (parliaments, elections), and media.

Media and Communication (ISSN: 2183–2439)
2021, Volume 9, Issue 4, Pages 5–15
https://doi.org/10.17645/mac.v9i4.4292

Article

Re-Defining Borders Online: Russia's Strategic Narrative on Internet Sovereignty

Anna Litvinenko

Institute for Media and Communication Studies, FU Berlin, Germany; E-Mail: anna.litvinenko@fu-berlin.de

Submitted: 21 March 2021 | Accepted: 4 August 2021 | Published: 21 October 2021

Abstract
Over the past decades, internet governance has developed in a tug-of-war between the democratic, transnational nature of the web, and attempts by national governments to put cyberspace under control. Recently, the idea of digital sovereignty has started to increasingly gain more supporters among nation states. This article is a case study on the Russian concept of a "sovereign internet." In 2019, the so-called law on sustainable internet marked a new milestone in the development of RuNet. Drawing on document analysis and expert interviews, I reconstruct Russia's strategic narrative on internet sovereignty and its evolution over time. I identify the main factors that have shaped the Russian concept of sovereignty, including domestic politics, the economy, international relations, and the historical trajectory of the Russian segment of the internet. The article places the Russian case in a global context and discusses the importance of strategic narratives of digital sovereignty for the future of internet governance.

Keywords
digital sovereignty; internet governance; Russia; strategic narrative

Issue
This article is part of the issue "Media Control Revisited: Challenges, Bottom-Up Resistance and Agency in the Digital Age" edited by Olga Dovbysh (University of Helsinki, Finland) and Esther Somfalvy (Research Centre for East European Studies at the University of Bremen, Germany).

1. Introduction

Global internet governance, which has been evolving over the past decades in the spirit of the Declaration of the Independence of Cyberspace (Barlow, 1996), has apparently reached a bifurcation point where more and more national governments are introducing their concepts of internet sovereignty. In the 2010s, this term had a rather negative connotation in the global media (Woodhams, 2019). The internet isolation policies of China, Iran, and Russia were seen as a destructive trend towards a "Splinternet," which would undermine the global digital economy and violate the human rights of freedom of speech and information access. When the so-called "sovereign internet" bill was introduced in Russia at the end of 2018, it was criticized in the press as an "online Iron Curtain" (Schulze, 2019). The widespread criticism and protests against the policy even made Russian legislators and pro-state media change the wording in the description from "sovereign internet" to "sustainable internet" (see Shimaev et al., 2019).

However, within the last few years, democratic countries such as EU states have also begun to talk intensely about their digital sovereignty (Pohle, 2020). Do different political regimes mean the same thing when they plead for digital sovereignty? Apparently not. The term "digital sovereignty" remains a highly contested one, and its interpretation differs from country to country and thus has "conflict potential" (Thiel, 2021). Kleinwächter (2021) has called the current state of Internet governance a "digital cacophony in a splintering cyberworld." This situation makes a study of strategic narratives (Miskimmon et al., 2013) of digital sovereignty an important contribution to the debate over the future of cyberspace.

In this article, I aim to explore Russia's strategic narrative regarding a "sovereign internet." Drawing on the

analysis of the major doctrines and strategies of the Russian government concerning internet policy since 1999, as well as on five expert interviews, I reconstruct the Russian government's strategic narrative of internet independence and explore the major factors that have shaped its approach to digital sovereignty. The article places the Russian case in a global context, contributing to a better understanding of the current challenges and perspectives of internet governance.

I conclude that the Russian concept is based on its approach to internet security, whereby internet security is likened to information security, and where the state's control over information flows is placed at the forefront. The key elements of the Russian understanding of digital sovereignty are: (1) control over data (in the form of data filtering and data localization), (2) control over infrastructure (in the form of, among others, protectionism and a centralized system of monitoring equipment), (3) promotion of Russian internet governance initiatives at the international level.

Although foreign threats to information security play a central role in Russia's strategic narrative of digital sovereignty, it is domestic politics and the impetus of elites to control oppositional discourse within the country that have apparently had the biggest impact on the formation of Russian digital sovereignty policy. I conclude by discussing the role of strategic narratives in regard to digital sovereignty for the future of Internet governance.

The remainder of this article is organized as follows: First, I give an overview of the major approaches to digital sovereignty. Then I present the development of Russian internet policy during the last decade, followed by the methodology section, the presentation of results and their interpretation, and the discussion.

2. Approaches to Internet Sovereignty: Drawing Borders in Cyberspace

On the one hand, the internet has opened immense opportunities for different actors worldwide. On the other hand, it has undermined the sovereignty of nation states, challenging existing rules and reshuffling the world order by empowering new global players, such as large social media platforms. In the 1990s and 2000s, the benefits of global interconnectedness for nation states largely prevailed over concerns about cyberthreats. However, after the Arab Spring in 2011, authoritarian leaders worldwide realized that the mobilizing potential of social media had become a real threat to their rule, so they have increasingly tightened control over online communication in their respective countries (Richter & Kozman, 2021). In 2013, Snowden's revelations about internet surveillance by US intelligence have stirred up a discussion about, among other factors, technical autonomy in EU countries (Müller, 2017).

While, in the 2000s, China with its Golden Shield program, also known as the Great Firewall, was a standalone example of internet isolation, in the 2010s, more and more countries started to follow this path. According to Mueller (2017), ideas about what we now call digital sovereignty were first introduced in different countries, including the US, long before 2013, but there has been no widespread rhetoric about the necessity of digital autonomy. Until recently, the term "digital sovereignty" has been associated mostly with authoritarian states, such as China and Iran.

These countries have, over the years, developed their national approaches to internet sovereignty. If we imagine a continuum, where openness of the net is on the right side and isolation on the left, the first from the left would be the case of North Korean, where the internet has been officially banned and replaced by a national intranet. The Chinese approach to internet sovereignty is much more sophisticated and apparently was able to solve the so-called "dictator's dilemma" (Kedezie, 1997). It implies that autocrats are usually faced with a choice between two paths that are both vital for the sustainability of their regime but, at the same time, contradict each other: the promotion of information technologies that bring economic benefits versus preserving control over the information space. The Chinese government manages to combine both these paths. It is, however, doubtful whether the Chinese case can be replicated, as the historical trajectory of internet development in China diverges from that of other countries. The internet in China was initially designed as a centralized network under state control. China's approach to internet sovereignty includes the Great Firewall, which filters undesirable content, and includes protectionism of Chinese IT companies and promotion of Chinese software and infrastructure worldwide (Steiner & Grzymek, 2020; Zeng et al., 2017). The Iranian approach is similar to the Chinese one, but it draws rather on a defensive strategy that was developed in reaction to international sanctions (Michaelsen, 2018). Exploring the factors that have shaped the current state of isolation of the Iranian internet, Michaelsen has highlighted the importance of international relations in this case.

Russia has joined the trend towards more state control over the internet rather late: The tightening of internet regulation there began after the protest movement "For Fair Elections" in 2011–2012 (Litvinenko & Toepfl, 2019). Russia's policy towards digital sovereignty has caused much discussion since the introduction of the 2018 draft of the bill on a sovereign internet, which was adopted in 2019 (Schulze, 2019).

For a long time, the EU has been rather reluctant to use the term "digital sovereignty" (Thiel, 2021), preferring the notions of "technical sovereignty" and autonomy (Pohle, 2020). Germany had a leading role in fueling the European debate on digital sovereignty by putting it on the agenda for EU digital policy during Germany's EU presidency in 2020 (Pohle, 2020). A year earlier, this term had been widely used in discussions about the project of the European data

cloud—Gaia-X—linking digital sovereignty to independence from externally produced infrastructure. German Minister of Economic Affairs and Energy Peter Altmaier said while introducing the project: "Germany has a claim to digital sovereignty. That's why it's important to us that cloud solutions are not just created in the U.S." (Stolton, 2019). In her speech at the opening of the Internet Governance Forum 2019 in Berlin, German Chancellor Angela Merkel gave the following definition of the concept: "Digital sovereignty does not mean protectionism, or that state authorities say what information can be disseminated—censorship….It describes the ability both of individuals and of society to shape the digital transformation in a self-determined way" (Merkel, 2019).

In her study of the European discourse on digital sovereignty, Julia Pohle has shown that, in the EU, the concept is linked to the democratic understanding of sovereignty as the people's right to self-determination (Pohle, 2020). It "encompasses the ability of individuals as well as state or commercial institutions to make autonomous use of digital technologies and to independently and securely exercise their roles in times of digitalization" (Pohle, 2020, p. 8). The existing definitions, however, are still too vague, as they need to be translated into tangible policy elements (Steiner & Grzymek, 2020).

So far, the term remains instead a metaphor that is interpreted in different ways by different political regimes. Kolozaridi and Muravyov (2021) have suggested understanding states' internet sovereignty claims as "performance, rhetorical acts whose primary function is to counter hegemonic tendencies." In a situation of a "digital cacophony in a splintering cyberworld" (Kleinwächter, 2021), the use of such a vague term might deepen existing controversies between states. At the same time, a better understanding of different states' narratives of internet sovereignty would bring more clarity to the ongoing processes of internet fragmentation.

Here, I suggest using the concept of strategic narrative that was shaped by Alister Miskimmon, Laura Roselle, and Ben O'Loughlin (Miskimmon et al., 2013). It is a theoretical framework for studying the persuasive communication of nation states in the international arena. By strategic narratives, they understand "a communicative tool through which political actors—usually elites—attempt to give determined meaning to the past, present, and future in order to achieve political objectives" (Miskimmon et al., 2013, p. 5). The authors distinguish strategic narratives at three levels: international system narratives, national narratives, and issue narratives (see also Roselle et al., 2014). The latter are meant to put governmental policies into context, and to explain why certain policies are necessary and how they can be successfully realized. Looking at the rationales that stand behind the use of the term "digital sovereignty" by different states will help us better understand the ongoing debate about the future of cyberspace.

3. The Russian Case: From an Underregulated Internet to Digital Sovereignty

The Russian segment of the internet, also called RuNet, remained largely unregulated until Putin's third presidential term, which started in 2012 (Vendil Pallin, 2017). In the 2000s, against the backdrop of increasing censorship in traditional media, the internet was celebrated as a free forum for political discussion (Richter, 2007). Scholars explained the absence of tight regulation over cyberspace by the fact "that the digital technologies do not offer a solution to issues of media control" (Richter, 2007, p. 206). However, as time has passed, new means to provide technological control over internet resources have emerged, which have been increasingly implemented by the Russian government.

The turning point in Russian internet policy was, according to many scholars, the protests of 2011–2012, which, to a large extent, were fueled by online media (Vendil Pallin, 2017). For instance, Soldatov (2015) mentioned that, although the blocking of websites had already been a rather common measure for the Russian authorities since 2007, it had previously been applied following a court decision and occurred in a non-systematic manner: "Since November 2012, internet censorship acquired a systemic nature" (Soldatov, 2015, p. 1).

In 2016, the so-called "Yarovaya Package," a set of amendments to anti-terrorism legislation, was adopted, which became an important milestone in the tightening of state control over cyberspace (Lehtisaari, 2019). Among other things, the law obliged internet providers to store all data for half a year, which was barely even technically possible. It also introduced more severe punishment for the (re)posting of pro-terrorist or extremist content.

One of the core characteristics of Russian internet legislation is its vague wording as well as its selectivity regarding the implementation of restrictive laws (Oates, 2013; Vendil Pallin, 2017). As Vendil Pallin has noted, "most laws are not systematically implemented and by no means all opposition content that is posted on the internet leads to legal or other actions from the authorities" (Vendil Pallin, 2017, p. 17).

Vendil Pallin (2017) examined strategies that the Russian government had implemented since 2013 in order to gain control over cyberspace though ownership of domestic resources and to regulate international companies operating on the RuNet—the first steps towards Russian digital sovereignty. For instance, in 2016, the obligation of internet operators to store the personal data of Russian citizens within the territory of the Russian Federation was officially framed "as a measure to increase security and safeguard the privacy of Russian internet users" (Vendil Pallin, 2017, p. 27). Another law that came into effect in November 2017 restricted the activities of VPN services and anonymizers, prescribing them to block Russian users' access to content prohibited by the Federal Service for Supervision in the

Sphere of Telecom, Information Technologies and Mass Communications (or Roskomnadzor). However, in the two years after the enactment of the law, VPN companies did not follow the rules, and the Russian authorities did not try to punish them for not doing so. As Soldatov mentioned in 2015, analyzing the perspectives of blocking the anonymizer Tor in Russia, just legal prohibition would be not enough, "a highly efficient technological solution is required" (Soldatov, 2015, p. 8), and it seems to not have been found yet.

The next major step on the path towards placing Russian internet segments under state control was the legislative initiative on "sovereign internet" that came into effect in October 2019. It introduced a system of state-sponsored monitoring devices that had to be installed by Internet providers and that helped authorities filter, reroute, and block internet traffic (Epifanova, 2020).

Stadnik (2021a) has analyzed internet independence policy in Russia by applying Müller's (2017) categorization of methods of alignment of cyberspace to national borders: national securitization, territorialization of information flows, and efforts to control critical internet resources along national lines. She has concluded that all these methods are being implemented in Russia and that the Russian government seeks to provide "national security at any price" (Stadnik, 2021a, p. 162), to a large extent ignoring the interests of private stakeholders. In her other paper, Stadnik (2021b) examined four attempts of the Russian government to exercise control over information flows via internet infrastructure, including a blacklist to filter internet content, the law on "sovereign RuNet," the failed attempt to ban the messenger app Telegram in the country, and a list of "socially significant websites" that could potentially be used as a "white list" of accessible internet resources. She concluded that these measures "do not work as the government would wish" (Stadnik, 2021b) and that content filtering leads to, among other things, undesirable side effects for the whole network.

Ramesh and colleagues did an investigative study of technical censorship mechanisms employed by the Russian state and came to the conclusion that the design of Russia's internet censorship in a decentralized network "is a blueprint, and perhaps a forewarning of what national censorship regimes could look like in many other countries" that have a network design similar to Russia's (Ramesh et al., 2020, p. 13). This makes the Russian case of internet control of significant importance for global internet governance, as it is potentially replicable in other countries, in contrast to that of China, where the internet is centralized by design.

After the introduction of the "sovereign internet" bill, several reports emerged analyzing Russian internet policy (Epifanova, 2020; Gruska, 2019; Soldatov, 2019). However, there is still a lack of academic research on the Russian approach to digital sovereignty. This study aims to address this gap by answering the following research question: What is the strategic narrative of the Russian government on internet sovereignty, and what are the main factors that have influenced its development?

4. Methodology

In order to address the research question, I have analyzed the official strategy papers on internet policies in Russia, issued by the government in the period 1999–2019, and I have conducted five semi-structured interviews with experts, which helped reconstruct the government's strategic narrative and identify the key factors in its evolution over time. The Russian official strategies and doctrines "feature the official position in regard to aims, tasks, principles and the main directions" of governmental policies (Russian Federation, 2016). They can thus be seen as an articulation of strategic narratives that are applied as fundamental principles for future legislation. In accordance with the terminology of Roselle et al. (2014), the narratives of official internet strategies in Russia can be categorized as issue narratives and are targeted both at the domestic audience to legitimize policies and at foreign governments as official messages in international politics.

From 1999 to 2020, the following seven strategic papers on internet policies were issued: Strategies of the Information Society Development (1999, 2008, 2017), Doctrines of Information Security (2000, 2016), Basic Principles for State Policy in the Field of International Information Security (2013), and Development Strategy for IT Industry for 2014–2020 and until 2025 (2013). I have also included the 2019 Federal Law 90-FZ, known as the "sovereign internet" bill, in the analysis, insofar as it contains a memorandum explaining the official rationale for introducing the bill.

The document analysis combined elements of content analysis and thematic analysis (Bowen, 2009, p. 32). It aimed at reconstructing the official state narrative in regard to independence in cyberspace by identifying the three elements in strategic narratives: problematized issues, claims of causality, and proposed solutions (Miskimmon et al., 2013; Szostek, 2017). In this particular case, it means focusing on the following categories: (1) key terms of internet policy, (2) rationales provided for policies in regard to independence in internet space, and (3) solutions—that is, policies themselves.

After completing the document analysis, I conducted five semi-structured interviews with experts on internet governance in Russia. The aim of the interviews was twofold: (1) to verify the findings of the document analysis and (2) distinguish the major factors that have led to changes in the strategic narrative on internet sovereignty over time.

The experts interviewed are representatives of different areas of expertise in Russian internet governance: Ilona Stadnik (Coordination Center for TLD .RU/.РФ), Michail Medrish (former head of the Coordination Center for TLD .RU/.РФ and member of the Council

of Europe's Committee of Experts on Cross-Border Flow of Internet Traffic and Internet Freedom), Andrei Soldatov (investigative journalist specializing in Russian internet policies), Polina Kolozaridi (researcher of the RuNet, associate professor at the National Research University Higher School of Economics), and Alena Epifanova (expert on Russia's domestic and foreign policy in cyberspace, German Council on Foreign Relations). In the interviews, I asked these experts to describe the Russian approach to internet sovereignty and how it differs from that of other countries, to name the milestones in the evolution of this approach, and the factors that, in their view, influenced this development. I also asked them to verify my conclusions from the document analysis. The interviews conducted via Skype were recorded, transcribed, and analyzed using NVivo software.

5. Findings

Below, I present the findings of the document analysis, according to the key areas of my inquiry: key terms of internet policy, rationales provided for policies in regard to independence on the internet, and solutions/policies.

5.1. Key Terms of Internet Policy

It is remarkable that the term "internet" is not mentioned in the strategic documents of 1999 and 2000 and is only mentioned three times in the eight pages of the 2008 Strategy of the Information Society Development. It was only in 2013 that the internet was mentioned prominently in the documents analyzed. The terms most widely used in all the documents are "information," "information sphere," and "information and communication technologies." In the 2000 Doctrine for Information Security, the role of the information sphere in the "strengthening of moral values of society" is emphasized. Here, for the first time, the necessity of "technological independence" for Russia in the IT sphere is mentioned.

The Information Society Development Strategy of 1999 sounds very optimistic and states that the main strategic goal of Russia in transition towards an information society is "the creation of a developed information and communication societal environment and Russia's integration into the global information community" (Russian Federation, 1999). In the 2008 strategy, the focus lay in the improvement of electronic governance, as well as in participation in international norm development and in the mechanisms of internet governance.

In both Doctrines for Information Security (2000, 2016), there is no mention of "cybersecurity," which is usually used in international documents. The focus is always on information, that is, on content, not on the channels of its transmission. According to these documents, Russia should counter "information threats," inter alia, information war. Although "information war" is an important term for the Doctrines, no clear definition of it is provided. Among the external threats to information security of the Russian Federation, "the development by a number of states of concepts of information wars" is listed, which implies "creation of means of dangerous impact on information spheres of other countries, violation of the normal functioning of information and telecommunication systems, safety of information resources, obtaining unauthorized access to them" (Russian Federation, 2016).

The 2017 Strategy of the Information Society Development places a bigger focus on the digital economy compared to those of 1999 and 2008. An important term in this strategy is "critical information infrastructure," which means information technologies used by state institutions and by different industries. In order to secure the critical information infrastructure, the state has to support and represent the interest of the national IT companies. In the 2017 document, one of the main aims of internet policy is a development in Russia towards being a "knowledge society," which is defined as a society "where the acquisition, preservation, production and dissemination of reliable information, while taking into account the strategic national priorities of the Russian Federation, are of predominant importance for the development of a citizen, the economy and the state" (Russian Federation, 2017).

In the 2013 Basic Principles for State Policy in the Field of International Information Security, "international information security" is defined as follows:

> A state of the global information space that excludes the possibility of violating the rights of an individual, society and the rights of the state in the information sphere, as well as destructive and illegal impact on the elements of the national critical information infrastructure. (Russian Federation, 2013a)

Following these principles, Russia should promote the establishment of an international legal order aimed at the "formation of an international information security system" (Russian Federation, 2013a).

5.2. Rationales and Solutions

The rationales behind the Russian internet policies in strategic papers have undergone a massive evolution over time. In the 1999 Strategy for Development of Information Society, the importance of preserving its independence in the process of globalization is mentioned, but the overall tone about globalization is optimistic and friendly towards the international community, which is even called a "family": "Russia has to join the family of technologically and economically developed countries as a full-fledged participant in the world civilizational development while maintaining political independence, national identity and cultural traditions" (Russian Federation, 1999).

According to this document, Russia has to find its own way in the information society, which would be

oriented to the Russian socio-cultural context and would require minimum financial investments from the state. This rather *laissez-faire* attitude of the state towards internet business is characteristic of the first decade of the 21st century.

In the 2000 Doctrine for Information Security, the list of the main threats to information security is not that long and is rather vaguely formulated: from threats to human rights to "the spiritual revival of Russia," to "information support of the state policy of the Russian Federation," and brain-drain of IT specialists. As a solution, the document emphasizes the importance of "information support of the policies of the Russian Federation," by providing the Russian and international audience with reliable information in this regard (Russian Federation, 2000). Support for Russian IT production is also highlighted.

The 2008 Strategy of the Information Society Development, which marked the start of the presidency of Dmitry Medvedev, was still optimistic towards information and telecommunication technologies, which, by then, had become "a locomotive of the socio-economic development" worldwide, so the state had to ensure "access of citizens to information" and develop e-governance services (Russian Federation, 2008).

The 2013 Strategy for Development of IT Industry for 2014–2020 mentions the increasing role of the internet in society: In 2012, the monthly audience of the internet in Russia reached more than 55 percent (Russian Federation, 2013b). "The absence of territorial borders on the internet" was seen as a chance for Russian IT companies to become leaders in the international market. Increasing the attractiveness of Russia as a jurisdiction for the operation of IT companies would positively affect the development of the domestic IT industry.

The 2013 Basic Principles for State Policy in the Field of International Information Security includes the promotion of Russian initiatives in the area of international information security. This is important, as ICTs can be used as, among other purposes, an "information weapon" for "discrediting sovereignty, violating the territorial integrity of states," and violating public order (Russian Federation, 2013a).

In the 2016 Doctrine for Information Security, the list of threats from ICTs become more articulate in comparison to the earlier 2000 document, and features, among other things, cybercrimes, terrorism, and the promotion of Russia-critical content by foreign actors. Moreover, according to the document, Russia runs a risk of being targeted by so-called "information weapons" due to "the intensive introduction of foreign information technologies" in Russian society. According to the 2016 Doctrine, these threats should be combatted by defending one's "own information sphere" from external influence. What exactly does this mean? For one thing, it means the so-called "import substitution" by national products and the protection of national interests in the market. Information security is to be provided not only by state authorities, but also by state media and telecom operators (Russian Federation, 2016).

The 2017 Strategy of the Information Society Development emphasizes the priority of "moral values traditional for Russia and social norms based on them when using technologies" (Russian Federation, 2017). This should be done by, for example, promoting information resources that are based on so-called "traditional Russian values." However, these values are not further defined.

The strategy papers starting from 2013 have increasingly mentioned various abstract foreign threats. The explanatory memorandum of the 2019 "sovereign internet" bill is more direct in its wording: It names the US as a threat to the sustainability of the internet in Russia. The bill was prepared "considering the aggressive tone of the US National Cyber Strategy adopted in September 2018" (Russian Federation, 2019). According to the memorandum, Russia was "groundlessly accused" by the US of commissioning hacker attacks and was threatened with punishment. The memorandum implies that this punishment could be the disruption of the country's internet. Therefore, according to the same document, in order to guarantee "a sustainable operation of the internet in Russia," preventive measures have to be taken. The bill implements technical means of countering "threats for integrity, sustainability and safety of functioning on the territory of the Russian Federation of the 'Internet' network" (Russian Federation, 2019). The so-called "sovereign internet" bill obliges internet providers to install devices provided by the state that can monitor and block internet traffic. These measures are thus presented as a preventive defense strategy against foreign threats.

5.3. Evolution of the Strategic Narrative on Internet Sovereignty Over Time

The analysis of strategic narratives on internet policies in official documents from 1999 to 2019 shows a shift that occurred around 2013: from perceiving the globalization of information primarily as a chance for, and source of, economic growth to focusing on threats that come with dependence on Western technologies and vulnerabilities of the open information space.

All documents emphasize that it is control over the content of information that matters first and foremost. According to the expert Andrei Soldatov, this constitutes a crucial difference from the Western approach to internet governance: "The Americans, the British, talked about cyber security, the security of wires, of power stations, that is, 'the iron.' And our officials have always used the term 'information security'… that is, content."

For Soldatov, the roots of the fundamental split in the understanding of the threat of the internet between Russia and the West lie in Russia's domestic affairs in the 1990s. According to the expert, at the beginning of the Second Chechen War in 1999, Putin had to explain

the government's failure in the First Chechen war, so he blamed it on the information interference of the journalists who covered the conflict. As a result, the Information Security Doctrine of 2000 stated the importance of the defense of the information sphere.

Based on the combination of the analysis of strategic narratives and expert interviews, the following elements of the Russian concept of internet sovereignty can be distinguished:

1. Control over data flows (i.e., filtering of content and data localization);
2. Control over infrastructure (i.e., protectionism of national software, centralized system of monitoring internet traffic);
3. Promotion of Russian initiatives at the international level.

The experts have distinguished the following factors, which, in their view, helped shape this approach: (1) domestic politics, (2) economic factors, (3) international relations, and (4) the historical trajectory of RuNet.

5.3.1. Domestic Politics

Inner political rationales were mentioned in the official documents and dealt with guaranteeing constitutional rights for citizens, as well as warranting stability, security, and economic progress in the country. However, as the experts confirmed, some of the important triggers for internet regulation were left out of sight in the official narrative. Thus, the protest movement "For Fair Elections" in 2011–2012 was crucial for the major shift towards the internet sovereignty that we observed in documents starting from 2013. The street protests broke out after the revelation of fraud during the parliament elections in December 2011 and demonstrated the power of social media in triggering an oppositional movement, which made the state reconsider its *laissez-faire* attitude towards regulation of online communication (Litvinenko & Toepfl, 2019). For the Russian government, said Alena Epifanova, this protest wave was apparently a more significant factor than the previous Arab Spring.

Another important trigger for the tightening of internet control was regional protests in 2017–2019. According to Andrei Soldatov, the blocking of the internet just in time in Russian regions in order to curb political dissent was one of the major aims of the "sovereign internet" bill.

Ilona Stadnik mentioned the case of the failure to block the Telegram messenger app in 2018–2020 as a catalyst for developing new mechanisms of control over the internet infrastructure. Telegram was officially blocked after the presidential election in 2018, but regulating institutions failed to stop its work. Citizens continued to use Telegram via VPNs, and it became even more popular, so the government decided to officially unblock it in July 2020.

Three experts also mentioned the role of elite power struggles within the Russian government, the so-called "war between the Kremlin towers." The first decade of internet development in Russia was dominated by more liberal elites, who were calling for Russia's modernization, especially under Medvedev's presidency in 2008–2012. Starting from the Putin's third presidential term in 2012, the role of *siloviki* (members of the ministries in charge of national security) has been increasing significantly. For them, security is more valuable than progress, and they tend to be in favor of internet blockages and other restrictions.

However, the government still cannot afford to simply cut Russia's access to global social media platforms, as it would most probably trigger major social unrest. Over the decades, people have gotten used to free communication online, and many users have built their businesses using the monetization models of YouTube or Instagram. According to Soldatov, this, among other factors, constitutes an important difference between the internet in Russia and in China. Thus, the government has to balance between its urge to control the information space and the risks of putting too much pressure on civil society.

5.3.2. Economic Factors

In the 2000s, the liberal approach to internet legislation was inspired by the perceived benefits of digitalization, which is reflected in the documents analyzed. An internet isolation policy, on the one hand, would mean losing many of those benefits. According to expert Ilona Stadnik, Russia cannot afford the risk of being disconnected from the global digital economy.

On the other hand, the aspiration of Russia to be independent in regard to internet infrastructure contradicts the current potential of the Russian IT industry. Despite the protectionism policy towards Russia's IT companies, Russia has no capacities to substitute all the imported IT products with Russian equivalents. Epifanova poses the question: "Will Russian Internet sovereignty be made in the US or in China?"

Over the past decade, Russian IT companies have been increasingly subjected to more control and compliance by the state. In 2016, the introduction of the Yarovaya law package, which obligated providers to store all communication data for six months at their own expense, stirred up a large protest within the IT industry. In 2019, the "sovereign internet" bill mandated that providers install equipment that would monitor internet traffic. However, Soldatov pointed out, this time the measure was to be paid for by the state, so the IT industry did not voice as much discontent as it had with the law of 2016. According to the expert, the IT companies realized that "with the current Russian image, they do not have many chances abroad anyway, so they have to develop the domestic market."

5.3.3. International Relations

This factor plays a crucial role in the state rationale behind the necessity of internet independence. Already in the 2000 Doctrine for Information Security, dependence on foreign IT companies is listed as one of the threats to national security. In 2019, the sovereign internet bill was framed as a reaction to the "aggressive tone" of the 2018 US National Cyber Strategy. The experts emphasized that the international relations factor was used rather as a tool to frame restrictive policies for the Russian audience.

The experts agreed upon the following international milestones, which had an influence on the Russian approach to internet sovereignty: (1) Edward Snowden's revelations in 2013, (2) international sanctions against Russia after the annexation of Crimea in 2014, and (3) the accusation of Russia in the interference in the US elections of 2016. Interestingly, as experts Kolozaridi and Stadnik point out, the Snowden revelations seem to have had less of an impact on Russian internet governance compared to the consequences they had in the West. In Russia, the digital sovereignty discourse started to evolve intensely after the introduction of economic sanctions in 2014 and the subsequent policy of import phaseout.

The scandal around the alleged interference of Russia in the US elections and around the data breach in Cambridge Analytica made the West reconsider its attitude toward cybersecurity. In the framework of fake-news debates, internet security is now also discussed in the West in terms of having control over content of information. Russia has perceived this as a window of opportunity to promote its understanding of information security, which it has already been sharing with China for a long time. According to Soldatov, legislation on fake news in different countries has given Russian authorities an opportunity to promote its narrative on information security.

Interestingly, dependence on global online platforms, which is central to digital sovereignty debates in the EU, has not been specifically thematized in the analyzed documents. However, this aspect has recently started to play a big role in public discourse and may be included in the strategic narrative on internet sovereignty in the future.

5.3.4. Historical Trajectory

In the interviews, all the experts mentioned the importance of the legacy of the historical development of RuNet, both from the technological and from the societal perspective, in shaping internet policies. According to Michail Medrish, the infrastructure of the Russian internet was initially designed to be highly decentralized, so it is hard to gain centralized control over RuNet. Soldatov elaborates that the liberal phase in internet regulation that lasted until 2012 shaped the country's online market as well as users' habits, and the state has been forced to take this into consideration on its path towards digital sovereignty. Epifanova concludes that the historical trajectory of RuNet makes the Russian model of digital sovereignty potentially attractive to other regions of the world, in contrast to the Chinese model, which is considered to be non-replicable.

6. Discussion and Conclusion

The Russian strategic narrative on internet policy has been changing over time, depending on the elite's evaluation of the benefits that global connectivity brings versus its perceived threats. The crucial element in Russia's understanding of internet independence is the concept of information security, which is content-oriented, in contrast to the Western concept of cybersecurity, which initially was more infrastructure-focused. This means that control over the content of data flows lies at the core of the Russian approach to internet sovereignty, and control over the infrastructure is seen as a tool to achieve this goal.

This contradicts the European understanding of digital sovereignty, which is based on the concept of the self-determination of the people (Pohle, 2020). As Europe has only recently coopted the concept, we are currently observing a global struggle of strategic narratives on digital sovereignty: a state-centered approach represented by Russia and China, where online borders are drawn maximally near the offline ones, and the individual-centered approach of the EU, where the people are called "sovereign."

However, the democratic interpretation of internet sovereignty appears, so far, to be even more vague than the authoritarian one, as democratic mechanisms of the self-determination of its netizens are still underdeveloped. Given the power of narratives in shaping the behavior of actors in international relations (Miskimmon et al., 2013), it seems to be important for international actors now to have an articulate vision and rationale for their approach to this widely used term. In a situation of struggle between strategic narratives around digital sovereignty, the promotion of a country's narrative at the international level becomes one of the key elements of internet sovereignty. This, among other things, helps build regional alliances among countries that hold similar positions on internet governance and thus gives more weight to arguments in favor of certain regulatory decisions on a global level.

The Russian case of internet sovereignty is of special importance for global internet governance, as it is an attempt to subject a highly decentralized network to tight state regulation via a series of measures, including control by infrastructure (Stadnik, 2021b). On the one hand, it has a direct influence on some post-Soviet countries where RuNet plays an important role, such as Belarus or Kazakhstan, and an indirect impact on many other segments of the global network through diffusion of legislative norms and practices. The global effects of

national approaches to digital sovereignty are still to be explored in future studies.

On the other hand, the Russian case reveals the weaknesses of the authoritarian model of digital sovereignty, which causes side effects for the network in the country. This model is also challenged by infrastructure-based resistance, as in the case of the attempted Telegram ban (Daucé & Musiani, 2021). Further study of the discrepancies between the norms and practices of digital sovereignty would help us better understand the mechanisms that shape today's internet governance.

Overall, we have observed that a strategic narrative on digital sovereignty is more than just an issue narrative, as it deals with a vision of the future of national segments of the internet as well as that of global internet governance. Comparative studies of national approaches to digital sovereignty are needed in order to define common ground for collaboration, as well as to distinguish between decisive divergences in envisioning the future of the global network.

Acknowledgments

This work was supported by the German Federal Ministry of Education and Research, funding code 16DII114, in the form of a fellowship of the Weizenbaum Institute for the Networked Society.

Conflict of Interests

The author declares no conflict of interests.

References

Barlow, J. P. (1996). A declaration of the independence of cyberspace. Electronic Frontier Foundation. https://www.eff.org/cyberspace-independence

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, *9*(2), 27–40. https://doi.org/10.3316/QRJ0902027

Daucé, F., & Musiani, F. (2021). Infrastructure-embedded control, circumvention and sovereignty in the Russian internet: An introduction. *First Monday*, *26*(5). https://doi.org/10.5210/fm.v26i5.11685

Epifanova, A. (2020). Deciphering Russia's "Sovereign internet law": Tightening control and accelerating the Splinternet. *DGAP Analysis*, *2020*(2). https://dgap.org/en/research/publications/deciphering-russias-sovereign-internet-law

Gruska, U. (2019). *Taking control? Internet censorship and surveillance in Russia*. Reporters Without Borders. www.reporter-ohne-grenzen.de/russiareport

Kedezie, C. R. (1997). *Communication and democracy: Coincident revolutions and the emergent dictator's dilemma*. RAND.

Kleinwächter, W. (2021, January 8). Internet governance outlook 2021: Digital cacaphony in a splintering cyberspace. *CircleID*. https://www.circleid.com/posts/20210108-internet-governance-outlook-2021-digital-cacaphony

Kolozaridi, P., & Muravyov, D. (2021). Contextualizing sovereignty: A critical review of competing explanations of the internet governance in the (so-called) Russian case. *First Monday*, *26*(5). https://doi.org/10.5210/fm.v26i5.11687

Lehtisaari, K. (2019). Formation of media policy in Russia: The case of the Iarovaia law. In M. Wijermars & K. Lehtisaari (Eds.), *Freedom of expression in Russia's new mediasphere* (pp. 57–73). Routledge.

Litvinenko, A., & Toepfl, F. (2019). The 'gardening' of an authoritarian public at large: How Russia's ruling elites transformed the country's media landscape after the 2011/12 protests 'For Fair Elections.' *Publizistik*, *64*(2), 225–240.

Merkel, A. (2019). *Speech opening the 14th Annual Meeting of the Internet Governance Forum in Berlin on 26 November 2019* [Speech transcript]. Press and Information Office of the Federal Government. https://www.bundeskanzlerin.de/bkin-en/news/speech-by-federal-chancellor-dr-angela-merkel-opening-the-14th-annual-meeting-of-the-internet-governance-forum-in-berlin-on-26-november-2019-1701494

Michaelsen, M. (2018). Transforming threats to power: The international politics of authoritarian internet control in Iran. *International Journal of Communication*, *12*(2018), 3856–3876.

Miskimmon, A., O'Loughlin, B., & Roselle, L. (2013). *Strategic narratives: Communication power and the new world order*. *Routledge studies in global information, politics and society* (Vol. 3). Routledge; Taylor & Francis Group.

Müller, M. L. (2017). *Will the internet fragment? Sovereignty, globalization, and cyberspace*. Polity.

Oates, S. (2013). *Revolution stalled*. Oxford University Press.

Pohle, J. (2020, December 15). Digital sovereignty: A new key concept of digital policy in Germany and Europe. *Konrad Adenauer Stiftung*. https://www.kas.de/en/single-title/-/content/digital-sovereignty

Ramesh, R., Raman, R. S., Bernhard, M., Ongkowijaya, V., Evdokimov, L., Edmundson, A., Sprecher, S., Ikram, M., & Ensafi, R. (2020). Decentralized control: A case study of Russia. In D. Xu & A.-R. Sadeghi (Eds.), *Proceedings 2020 network and distributed system security symposium* (pp. 1–18). Internet Society. https://doi.org/10.14722/ndss.2020.23098

Richter, A. (2007). *Post-Soviet perspective on censorship and freedom of the media*. UNESCO Moscow Office.

Richter, C., & Kozman, C. (Eds.). (2021). *Arab media systems*. Open Book Publishers. https://www.openbookpublishers.com/product/1281

Roselle, L., Miskimmon, A., & O'Loughlin, B. (2014). Strategic narrative: A new means to understand soft power. *Media, War & Conflict*, *7*(1), 70–84. https://doi.org/10.1177/1750635213516696

Russian Federation. (1999). *Kontseptsiya formirovaniya informatsionnogo obshchestva v Rossii* [Concept of development of information society] (No. 32). http://emag.iis.ru/arc/infosoc/emag.nsf/BPA/37cd5e6756dce634c32568c000474a8a

Russian Federation. (2000). *Doktrina informatsionnoy bezopasnosti Rossiyskoy Federatsi* [Doctrine of information security of the Russian Federation] (N Pr-1895). http://base.garant.ru/182535

Russian Federation. (2008). *Strategiya razvitiya informatsionnogo obshchestva v Rossiyskoy Federatsii* [Strategy of the information society development in the Russian Federation] (N Pr-212). https://rg.ru/2008/02/16/informacia-strategia-dok.html

Russian Federation. (2016). *Doktrina informatsionnoy bezopasnosti Rossiyskoy Federatsi* [Doctrine of information security of the Russian Federation] (N 646). http://ivo.garant.ru/proxy/share?data=q4Og0aLnpN5Pvp_qlYqzjK_xqrzXt9W_qeqZArb1tcalo_yf8-aowbnJtcvygADzs-CA4ZPhgf2P5pb8nfPvvualzLXQpdG50wLqneeE5LXnseO8rQ

Russian Federation. (2013a). *Osnovy gosudarstvennoy politiki Rossiyskoy Federatsii v oblasti mezhdunarodnoy informatsionnoy bezopasnosti na period do 2020 goda* [Basic principles for state policy in the field of international information security to 2020] (N Pr-1753). http://www.consultant.ru/document/cons_doc_LAW_178634

Russian Federation. (2013b). *Strategiya razvitiya otrasli informatsionnykh tekhnologiy v Rossiyskoy Federatsii na 2014–2020 gody i na perspektivu do 2025 goda* [Strategy for the development of the information technology industry in the Russian Federation for 2014–2020 and for the future until 2025]. https://digital.gov.ru/ru/documents/4084

Russian Federation. (2017). *O strategii razvitiya informatsionnogo obshchestva v Rossiyskoy Federatsii na 2017–2030 gody* [On the strategy for the development of the information society in the Russian Federation for 2017–2030] (No. 203). http://publication.pravo.gov.ru/Document/View/0001201705100002

Russian Federation. (2019). *Zakonoproyekt № 608767-7 O vnesenii izmeneniy v Federal'nyy zakon «O svyazi» i Federal'nyy zakon «Ob informatsii, informatsionnykh tekhnologiyakh i o zashchite informatsii"* [Draft Law No. 608767-7 on amendments to the federal law on changes in the federal law "On communication" and the federal law "On information, information technologies and information protection"]. https://sozd.duma.gov.ru/bill/608767-7

Schulze, E. (2019, November 1). Russia just brought in a law to try to disconnect its internet from the rest of the world. *CNBC*. https://www.cnbc.com/2019/11/01/russia-controversial-sovereign-internet-law-goes-into-force.html

Shimaev, R., Peletayeva, P., & Rumyanzeva, A. (2019, April 16). "Otklyuchit' rubil'nik uzhe ne poluchitsya": Gosduma utverdila zakon o bezopasnom i ustoychivom internete ["Turning off the switch will no longer work": The State Duma approved the law on a safe and sustainable internet]. *Russia Today*. https://ru.rt.com/dbxj

Soldatov, A. (2015). Ukroshtshenie interneta [Taming the internet]. *Contrapunkt*, *2015*(1), 1–11.

Soldatov, A. (2019). Security first, technology second: Putin tightens his grip on Russia's internet—With China's help. *DGAP Policy Brief*, *2019*(3). https://dgap.org/en/research/publications/security-first-technology-second

Stadnik, I. (2021a). Russia: An independent and sovereign internet? In B. Haggart, N. Tusikov, & J. A. Scholte (Eds.), *Power and authority in internet governance: Return of the state?* (1st ed., pp. 147–167). Routledge. https://doi.org/10.4324/9781003008309

Stadnik, I. (2021b). Control by infrastructure: Political ambitions meet technical implementations in RuNet. *First Monday*, *26*(5). https://doi.org/10.5210/fm.v26i5.11693

Steiner, F., & Grzymek, V. (2020). *Digital sovereignty in the EU*. Bertelsmann Foundation. https://www.bertelsmann-stiftung.de/en/publications/publication/did/digital-sovereignty-in-the-eu-en

Stolton, S. (2019, September 12). Altmaier's cloud initiative and the pursuit of European digital sovereignty. *Euractiv*. https://www.euractiv.com/section/data-protection/news/altmaiers-cloud-initiative-and-the-pursuit-of-european-digital-sovereignty

Szostek, J. (2017). The power and limits of Russia's strategic narrative in Ukraine: The role of linkage. *Perspectives on Politics*, *15*(2), 379–395. https://doi.org/10.1017/S153759271700007X

Thiel, T. (2021, January 25). Das Problem mit der digitalen Souveränität [The problem with digital sovereignty]. *Frankfurter Allgemeine Zeitung*. https://zeitung.faz.net/faz/unternehmen/2021-01-25/4b6c5ef358b56fe3c17d9912315df988/?GEPC=s3

Vendil Pallin, C. (2017). Internet control through ownership: The case of Russia. *Post-Soviet Affairs*, *33*(1), 16–33.

Woodhams, S. (2019, April 23). The rise of internet sovereignty and the end of the world wide web? *The Global Post*. https://theglobepost.com/2019/04/23/internet-sovereignty

Zeng, J., Stevens, T., & Chen, Y. (2017). China's solution to global cyber governance: Unpacking the domestic discourse of "internet sovereignty." *Politics & Policy*, *45*(3), 432–464. https://doi.org/10.1111/polp.12202

About the Author

Anna Litvinenko (PhD) is a researcher in the Digitalization and Participation Department at the Institute for Media and Communication Studies, FU Berlin, Germany. After receiving her PhD in 2007, she was associate professor in the Department of International Journalism at St. Petersburg University, Russia. Her research focuses on political communication in the digital age, comparative media studies, and the role of social media in various socio-political contexts.

COGITATIO

Media and Communication (ISSN: 2183–2439)
2021, Volume 9, Issue 4, Pages 16–26
https://doi.org/10.17645/mac.v9i4.4177

Article

From "Troll Factories" to "Littering the Information Space": Control Strategies Over the Russian Internet

Ilya Kiriya

School of Media, HSE University, Russia; E-Mail: ikiria@hse.ru

Submitted: 12 February 2021 | Accepted: 17 June 2021 | Published: 21 October 2021

Abstract
This article explores aspects, transformations, and dynamics of the ideological control of the internet in Russia. It analyses the strategies of actors across the Russian online space which contribute to this state-driven ideological control. The tightening of legislative regulation over the last 10 years to control social media and digital self-expression in Russia is relatively well studied. However, there is a lack of research on how the control of the internet works at a structural level. Namely, how it isolates "echo chambers" of oppositional discourses while also creating a massive flood of pro-state information and opinions. This article argues that the strategy of the Russian state to control the internet over the last 10 years has changed considerably. From creating troll factories and bots to distort communication in social media, the state is progressively moving towards a strategy of creating a huge state-oriented information flood to "litter" online space. Such a strategy relies on the generation of news resources which attract large volumes of traffic, which leads to such "trash information" dominating the internet.

Keywords
alternative media; ideological control; digital self-expression; power; RuNet; Russian media; social media

Issue
This article is part of the issue "Media Control Revisited: Challenges, Bottom-Up Resistance, and Agency in the Digital Age" edited by Olga Dovbysh (University of Helsinki, Finland) and Esther Somfalvy (Research Centre for East European Studies at the University of Bremen, Germany).

1. Introduction

The question of ideological control over the Russian information space remains central in understanding the peculiarity of Russian, and even more broadly, post-Soviet media systems. The institutional side of the post-Soviet media system has been studied relatively well, especially the role of "oligarchs" (Mickiewicz, 2008; Zassoursky, 2016) in distorting the democratic model of media. Macro-analysis of the whole media system has been conducted quite thoroughly (Kiriya, 2019; Oates, 2007; Vartanova, 2011). Other researchers have preferred the micro-social approach and have demonstrated how ideology is created based on routinised actions such as self-censorship (Koltsova, 2006; Schimpfossl & Yablokov, 2014).

The topics of the internet and mass-self communication are often not included in the analysis of the dynamics of the post-Soviet media system and ideological control. Prior to 2011–2012, the internet was not a topic considered in the analysis of post-Soviet media. After the Moscow 2011–2012 uprisings (i.e., Moscow uprisings of winter 2011 and spring 2012, provoked by the movement For Fair Elections), the topic of the so-called new media's role in post-Soviet social dynamics became more visible in media studies. Conversely, although the role the internet plays in the overall media system has been idealised for a long time, the idea that the internet represents a kind of new liberal force or alternative media in which the agenda differs drastically from traditional media has since been refuted. Thus, most research now focuses on new hyper-restrictive legislation bringing the RuNet under control of the Kremlin (Gabdulhakov, 2020), analysis of agenda and topical clusterisation of the Russian discourse in social media (Koltsova & Shcherbak, 2015), the role of the Russian trolls and hackers in the 2016

US presidential elections (McCombie et al., 2020), and so forth. All visions of scholars can be divided into two polar camps: internet-optimism (expressing hopes that internet discourse is creating the alternative public which will positively contribute to liberal dynamics of the media development) or internet-pessimism (generally arguing that the internet is just contributing to the isolation of oppositional groups, their marginalisation, and formation of echo-chambers).

The techno-deterministic vision of the internet as a new actor opposing the "old media order" has been dominant and has finally led to the relative autonomy of the internet from other media systems in analytical frameworks. Until now, we have usually read forecasts about the huge difference between the RuNet audience and the television audience. Such differences between these two audiences were central in the denomination of protest groups during the uprisings of 2011–2012 and 2019 as "hipsters" against "vatniks."

Nobody has tried to understand, from a systematic point of view, the role of the internet in state communication control and inside the whole media system, including the still powerful "traditional" television and periodical press sectors. In our own approach, we prefer not to compare the internet to traditional media, their audience, or content, but instead look at them as dynamic systems of information control oriented towards maintaining the dominant order, the high level of trust in the president, and core state institutions. In this article, we will trace the main changes in state control over the internet in Russia over the last 10 years and show how this control works together with other institutional mechanisms, to ensure the restrictive and state-oriented character of the whole media system.

Our basic hypothesis counters the idea of the liberalising mechanism of the internet for Russian society and shows that the state has progressively changed its methods of control over the internet according to changes in general media consumption. This includes methods such as "troll factories," progressively passed towards other more structural incitements related to the massification of internet use. In this conclusion, we are generally adding to the idea that the Russian model of restricting mass media is rather different from direct censorship based on filtering and technical blocking, as is the case in China and Iran (Toepfl, 2018, p. 542).

2. (De)Mythologising the Internet as Alternative Media

Before 2011, the internet did not appear among major topics discussed on the Russian media system, while the academic analysis of this field was primarily focused on the renaissance of Soviet rules of journalism (Oates, 2007) or the emergence of a "neo-authoritarian media system" (Becker, 2004) and the peculiarity of post-Soviet media systems (Vartanova, 2011). Thus, Vartanova (2011) clearly compares the emerging internet (at the time of writing in 2010, it was not so massively used by its audience) to the remaining part of the media sphere, saying that "marginalized forces opposing state influence in the media (investigative and opposition journalists, internet activists, and active audiences) have been active in promoting a free press, a free internet, and ethical norms in new media" (p. 142).

Thus, for a long time, the idea of the internet as a tool of resistance to the conservatism and state dirigisme of the Russian media landscape was mainstream in studies on the Russian media and internet. In 2010, the Berkman Centre at Harvard University made the map of the Russian blogosphere, where the internet was represented as an alternative public discussion arena where liberal opposition could coexist with other marginal political movements outside the mainstream spectrum (Etling et al., 2010).

Bode and Makarychev (2013) argued that the potential of the new social media was substantial, especially when compared with the Kremlin's loss of ability to generate socially acceptable meanings, "to convey messages to target audiences, to dominate the symbolic and ideational landscapes, and ultimately to maintain its discursive hegemony" (p. 61). Koltsova and Shcherbak's (2015) study of the online discourse of the 2011 uprisings shows that "the blogosphere belonged predominantly to oppositional bloggers" (p. 1724). In conclusion, they pointed out that in comparison with over-censored TV and mainstream media, the internet represented an arena for alternative political communication (Koltsova & Shcherbak, 2015, p. 1727).

What is quite emblematic is that even after the Crimean consensus (when a large part of the population welcomed the state's geopolitical game), the positive vision of the internet as enabling opposition with real political power continued to exist inside academia. Thus, Remmer (2017) argued that the internet "facilitated the formation of personal networks of digital activists who challenged the regime's control of the public sphere and offered an alternative discourse to the official political narrative" (p. 126).

As we can see in all such approaches, the internet and social media are especially associated with some holistic entities opposing the mainstream discourse. Even if all previously mentioned authors never used an alternative media framework to represent the subversive potential of RuNet, the opposition they have established between internet and non-internet media agendas pushes us to examine the alternative media concept from the Russian media landscape perspective.

The idea of alternative media has been well formulated by Bailey et al. (2008, p. 6) as based on four different approaches: (1) alternative media as serving the community; (2) alternative media as opposing the mainstream media; (3) alternative media as serving civil society; and (4) alternative media as a rhizomatic concept (emphasising the purely floating sense of the term).

The first approach cannot be directly applied to the Russian internet because the internet does not serve

a particular community. Social media on the internet can contribute to the creation of communities, but even in the case of the Coordination Council of Opposition formed just after the Moscow uprisings of 2011, it is quite difficult to call this a community because of its very heterogeneous and strong participatory nature. The second approach can be applied only if we understand perfectly what we mean by "mainstream media." However, not all principles of alternativity can be applied to the RuNet. Bailey et al. (2008, p. 18) give four characteristics of mainstream media: (1) large-scale and geared towards large, homogeneous (segments of) audiences; (2) state-owned organisations or commercial companies; (3) vertically (or hierarchically) structured organisations staffed by professionals; and (4) carriers of dominant discourses and representations.

When one compares RuNet to the mainstream media, it is represented as: (1) small-scale and oriented towards fragmented audiences; (2) non-controlled, either by the state or commercial bodies; (3) horizontally structured and run only by non-professional politicians (citizens themselves); and finally, (4) opposing dominant discourse and representations. We argue that the Russian internet does not match all these criteria. First, it is not always oriented towards fragmented audiences or low scale: large mainstream media corporations are active within it. Second, politically opposing content on the internet is not always created outside the commercial realm or state control, and even social media represents predominantly commercial corporations who earn money from the users' activity. Third, not all content of political opposition in social media is created horizontally. Some of it is organisationally enabled, and after the 2011 uprisings, the level of organisational control of such activities became even higher with the election of the Coordination Council of the Opposition (Toepfl, 2018). Finally, not all internet and social media oppose the dominant discourse. This argument could also be considered "universal" and fair for all other media landscapes and not only Russian ones. We can see generally in the wider world that not all internet media can be considered "alternative" since a large part of internet audiences are generated by organisationally enabled commercial media, which use the internet as a new way to produce surplus value and maximise profit (Fuchs, 2014).

At the same time, the Russian peculiarity is that the borderline between so-called grassroots media and elite or organisationally backed media (such distinction is based on Fuchs, 2010) is blurred. We know of some examples when media initially created by a group of independent journalists (sometimes a group of journalists who had been fired from big media for political reasons) rapidly gained some powerful investors. This is, for example, the case of Meduza.io, the internet media created and based in Riga by a self-organized group of journalists fired from Lenta.ru. This creation of Meduza could be considered a grassroots initiative of a group of journalists. However, we know (Surganova, 2014) that Meduza is financed by some undisclosed oligarch and that its founder, ex editor-in-chief of Lenta.ru, Galina Timchenko, negotiated financial issues with Michail Khodorkovsky (a Russian oligarch in exile in London). From this point of view, we might consider Meduza a classic commercial dependent media. In the case of some Russian offline media such as TV Rain or, for example, *Novaya Gazeta*, such distinction might also be problematic. On one hand, such media self-position themselves as a community of critically thinking journalists, and they also rely on grassroots business models such as crowdsourcing. *Novaya Gazeta* proposes that readers "support the independent journalism by making donations" in the disclaimer at the end of each publication. TV Rain is subscription-based, but this television channel communicates with subscribers as contributors whilst at the same time being privately owned. *Novaya Gazeta* is co-invested by Aleksander Lebedev, a Russian liberal oligarch, very well-known in elite circles and an ex-officer of the Russian KGB.

The third approach (alternative media as civil society media) does not work either, because the oppositional forces in Russian social media are very heterogeneous and do not necessarily rely on civil society structures. Some of them act on behalf of wealthy oligarchs (such as internet-media MBH-Media, owned and financed by Michael Khodorkovsky) or other elite-based structures. Finally, the rhizomatic approach to alternative media is also deficient in the case of RuNet because, as Bailey et al. (2008) argue, alternative media plays the catalytic role in "functioning as the crossroads where people from different types of movements and struggles meet and collaborate." As Kiriya (2012, p. 461) argued, RuNet is much more oriented towards isolating and fragmenting communities rather than uniting them.

The Russian internet is multi-level and multi-faceted and should simultaneously be considered as a means of resistance and a means of maintaining ideological order. For a long time, academic discourse has privileged its resistant side without seriously considering its ability to control and maintain the dominant order. Such discourse can be explained. In 2010, big state-owned and oligarchically supported media never considered the internet as an important source of audience and revenue, while some oppositional media outlets, such as self-organised media, considered the internet as a kind of parallel public sphere with a more intellectual audience, more oriented towards a Western way of life and civic freedoms.

3. The Internet as a Part of the Controlled and Surveilled Media Sphere in Russia

In parallel to the mainstream media studies' discourse on the resistant character of the RuNet, we can see some studies appearing in the second decade of the 2000s trying at least to question the emancipating character of the internet. Deibert and Rohozinski (2010) made quite a full review of different methods of internet control on

post-Soviet space more than 10 years ago, although, at that time, Russian state control over the internet had not yet even begun. However, we will largely use their framework in the next part when analysing the dynamics of state control.

The most popular way to incorporate new media within the whole Russian media system was a "fragmentationist" approach, which tried to represent Russian media as a set of a few public spheres with different rules with the state merely acting as a gatekeeper between them. Toepfl, in the early 2010s, showed how Russian ruling elites managed public scandals originating from "new media postings" by providing them with biased coverage within mainstream media and re-framing them in a manner that did not harm the dominant regime's legitimacy (Toepfl, 2011). Kiriya (2012) showed the fragmentation of the Russian public sphere which contributed to the maintenance of the relative pro-state order in RuNet. Bodrunova and Litvinenko (2015) analyse the fragmentation of the Russian public sphere mainly on the basis of the fragmentation of the population.

The isolation of oppositional communities and their concentration around oppositional media allows the government to "promote dominant agenda via state-controlled outlets" and to monitor protest moods by surveilling such oppositional information ghettos (Denisova, 2017, p. 989). Moreover, the participatory content created and shared within such a community could be regarded as an alternative to protest mobilisations and has even been tolerated by the state (Karatzogianni et al., 2017, p. 120). We should be very careful in stressing that such oppositional ghettos are organised around the internet media. As Oates (2016) wrote: "There was no complete division of the public between anti-Putin/online and pro-Putin/traditional media….This underlines the point that the internet is not a sphere separate from the political and media logics of the state" (p. 410).

Other studies describing internet control in Russia have been rather oriented towards analysis of measures implemented by the Russian government and parliament to place RuNet under their control. Here we can mention works analysing 2013 anti-piracy laws (Kiriya & Sherstoboeva, 2015), online self-expression regulations (Gabdulhakov, 2020), the corporate takeover of internet companies (Vendil Pallin, 2017), and implicating users and volunteers through surveillance and control of internet content (Daucé et al., 2019).

More generally, there is a lack of work researching the general philosophy of the Russian state towards the internet. Budnitsky and Jia's (2018) analysis of the Russian and Chinese policies in the field of internet sovereignty represents a good overview of this field.

In this article, we are trying to put all these methods of control together to show the dynamics and understand the strategy of the state in this field in its complexity. For the analysis, we rely on models of control distinguished by Deibert and Rohozinski (2010) and on the theory of alternative media. In our opinion, the core shift in state regulation of the internet is related to mainstream/alternative cleavages. Thus, the core hypotheses of this article are:

> H1: The core difference of the Russian model of networked authoritarianism is in balancing between the open prohibition and structural measures affecting the circulation of messages between "mainstream" and alternative media;
>
> H2: The usage of the internet inside the mechanism of control of the whole Russian media system has evolved considerably over the last 10 years and has depended hugely on the massification of the internet as a communicative platform.

4. Dynamics of Russian Internet Control

4.1. Between Censorship and Self-Censorship

Despite the digital pessimism of some analysts who interpreted the tightening of internet regulation in Russia as a step towards the building of a great firewall (Roth, 2019), the "Chinisation" of the Russian internet still seems far off. For instance, the Russian strategy seems considerably different from the Chinese, according to Toepfl (2018, p. 542), at least three elements contribute to their difference: (1) In contrast to China, the mass media landscape in Russia is only partly controlled by authoritarian elites—the alternative opinion space is shrinking, but now the public can have access to alternative partisanship media, foreign media, and media financed by foreign institutions; (2) unlike in China, the Russian internet as a communicative space is not subject to large-scale technological filtering—even having adopted some restrictive laws, the control is more post-publication based rather than the real filtering of prior publications; and (3) opposition groups, NGOs, and parties can operate legally in Russia, even if the state puts them under considerable control.

Deibert and Rohozinski (2010) distinguish three generations of internet control in the post-Soviet space. The first generation is based on denying access to specific online media by directly blocking access to servers, domains, keywords, and IP addresses. The second aims to create a legal framework for denying access and considerably reducing the possibility of discovering certain content.

The third generation is based on proposing competing content and making counterinformation campaigns to discredit opponents. When this framework was proposed, the Russian internet remained a relatively free space, and many of these measures were quite far away. However, since the Moscow uprisings, we can observe some dynamics in the tightening of control over RuNet. It is obvious that Russian authorities started to put more effort into restricting the internet after the Moscow

uprisings of 2011–2012 because social media played some role in the mobilisation movement. At the same time, the core reason for changing the policy of Russian authorities on the internet should be related to crucial changes in the configuration of the public sphere(s) since then.

The core change is related to the massification of the internet as an information space. In 2011–2012, the share of monthly internet users in Russia was 46%. If we take the most active daily users, this figure was 33% (Public Opinion Foundation, 2011). As usual, not all used the internet as a news source. As a result, during this period, the politically active internet audience was not so significant (around 10% of the population) and was essentially concentrated in big cities where generally the level of opposition votes is higher. At the end of 2020, the share of monthly users reached 78% and daily users reached 71% (Mediascop, 2021). From this point of view, we can understand why some researchers apprehended the internet as an alternative source of information in 2011–2012 because it was a platform for a relatively prosperous minority. Therefore, the Moscow uprisings were called the hipsters' protests or the creative-class uprisings (Goryashko & Prokofiev, 2012). However, in 2020, such characteristics are problematic due to the huge massification of the internet. Today 88% of internet users report having consumed television at least once in the past two weeks, a figure which has been in a slight decline over the last five years (Figure 1).

It is obvious that such quantitative changes in the number of internet users cannot be ignored by state policy in the field of the internet. The initial practice largely used against opposition leaders online was trolling, with the introduction of "trolls" into the debate distorting sustainable communication. Such pointed measures were probably enough to marginalise and distort opposition minorities online, but since 2013, the state has operated a massive campaign by creating a legal restricting framework oriented towards legitimising the blocking and the denial of access to internet sites. Here we can refer to the anti-piracy law, which can be used to block some resources (Kiriya & Sherstoboeva, 2015). Similar cases are related to progressive criminalisation of the users' activity on social media (such as Article 148 in the Criminal Code "in the aim of protecting religious convictions and feelings"; Article 205 on the endorsement of terrorism, etc.; Russian Federation, 2021). Finally, we can mention the 2019 law on "fake news" giving unprecedented rights to block content considered fake news, as well as any content deemed to be insulting to the authorities (Russian Federation, 2019b). From this point of view, we can stress that Russia adopted the second strategy of the Deibert-Rohozinski model (Deibert & Rohozinski, 2010) without employing the first one.

Another strategy that we can partly classify as the second model of control is oriented towards the development of self-censorship and is based on two tools. The first is creating fear among internet users that any action inside the mass-self communication can be tied to repressive legislation. A good example of this is Law 530 FL "On information, information technologies, and information protection" from 30 December 2020, which since 1 February 2021 obliges social media to block any obscene words in users' posts without really explaining how the law will be applied and without introducing clear responsibility (Russian Federation, 2020). Since technically blocking all obscenities seems to be impossible, the authorities are able to apply the law selectively. The second tool is relying on collective moral and taste norms, which are massively introduced online by internet user associations within the framework of "digital vigilantism," so a kind of parallel to the police form of civic

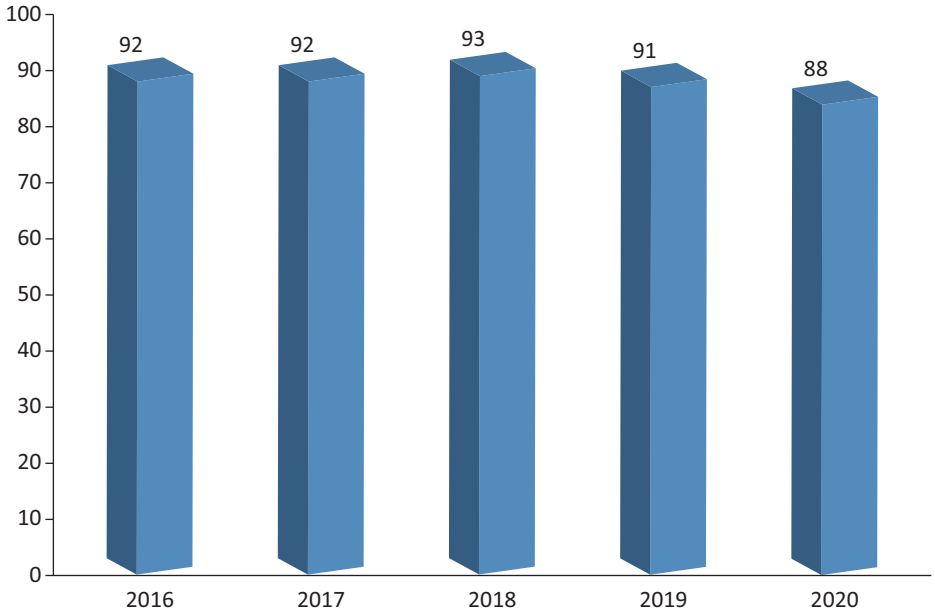

Figure 1. Share of internet users who watched TV at least once during the last two weeks. Source: Deloitte (2020).

enforcement: cyber patrols, cyber cossacks, leagues, and some other organisations oriented towards surveillance of users (Daucé et al., 2019).

Together, such tools show that Russian police and security authorities are unable to curb the resistant potential of new media without relying on "superficial measures designed to stimulate self-censorship" (Gabdulhakov, 2020, p. 297). They are creating the surveillant assemblage, which to a great extent works without enforcement from the state, but is based on other disciplining practices from various other actors which influence the mass behaviour of users (Gabdulhakov, 2020).

To simplify its task of blocking and controlling the parallel realm of social media, the state uses the same strategy as in the field of traditional media. It acquires the capital of major social media platforms under the financial control of loyal oligarchs. Thus, close to the Kremlin oligarch, Alisher Usmanov and his Mail.ru Group obtained control of VK (the most popular social media platform) by putting pressure on its previous owner and founder, Pavel Durov (Vendil Pallin, 2017, p. 25). While the president of Mail.ru Group is Boris Dobrodeev, very close to pro-power elite circles.

In parallel with building surveillance practices, the state maintains the public interest in such issues as internet sovereignty. The adoption of the law on "sovereign internet" (Russian Federation, 2019a), which gives the ability to organise internet traffic and routing locally in case of exterior disconnection, plays the role of the public trigger for any forces relying on local regulation of content and self-censorship. The law, together with massive coverage of the Russian hackers' infringement into US elections, creates among mass users the feeling that we are dealing with a cyberwar.

All mechanisms of self-censorship show that the mechanism of internet control in Russia is balancing between direct oppressive measures and the creation of an atmosphere disciplining users themselves or orienting them towards pro-state behaviour.

4.2. Between Mainstream and Alternative

Kiriya (2012) proposed the framework of the "parallel public sphere" to interpret the diversity of agendas between the internet and mainstream media. The public sphere has been divided into the mainstream public sphere and the parallel public sphere. The parallel public sphere, in turn, has been divided into a parallel institutionalised public sphere and a parallel non-institutionalised public sphere. The parallel institutionalised public sphere represented some official media (both offline and online) existing as organisations and oriented towards opposition points of view. It includes *Novaya Gazeta*, Echo Moskvy radio station, Ren-TV channel, and TV Rain, amongst others. A huge part of such opposition outlets was under the financial control of big oligarchic groups (such as Gazprom, state loyal bankers, etc.). In recent years, we may include the MBH media internet portal financed by Michail Khodorkovsky (was closed in 2021), Meduza.io portal based in Riga. Some media became less oppositional and much more pro-Kremlin (such as Ren-TV channel). The non-institutionalised parallel public sphere was represented by grassroots projects existing only based on social media, video sharing, and blogs. The core difference between institutionalised and non-institutionalised media was the greater level of pressure on the institutionalised parallel public sphere. Since the internet audience in this period was much more different from the mainstream media audience and pretended to have a much higher degree of partisanship, even the mainstream media on the internet proposed a more liberal agenda than their purely offline colleagues. For example, the state-financed information agency Ria Novosti covered the Moscow uprisings quite broadly (which became one of the reasons the editor-in-chief of Ria Novosti, Svetlana Mironyuk, was forced to resign by the Kremlin). From this point of view, when the internet was relatively young and assembled the oppositional public, there was a broad phenomenon of "alternativisation" of mainstream media.

Massification of the internet, including online news consumption, means much broader audiences are unlikely to remain free from the states' attention, as it started to increase its symbolic presence within the net. From this point of view, in parallel to the strategy of prohibitions, penalisation, and restriction of self-expression in social media, the state starts to build a system countering oppositional messages which we might call the "mainstreamisation" of the alternative media.

To show the level of state control over internet resources and its evolution between 2012 and 2020, the author analysed and classified the main internet outlets of these two periods of time based on their political orientation. The author took the main media internet outlets in 2012 according to data of TNS (Russian audience measurement company of this time, projection Web Index) based on their audience (average issue readership). Later, new news media outlets that appeared after 2012 were added to this list.

Some methodological remarks must be made prior to analysis. As pointed out by Degtereva and Kiriya (2010), there are three types of state control over the media on the level of ownership: (1) media directly owned by the state (e.g., Channel One where 51% of shares are directly owned by the Russian government); (2) media owned by the state, but may have some private monetary capital (e.g., NTV channel owned by Gazprom); and (3) media owned by state-loyal oligarchs related by non-formal connections with ruling elite groups and the president personally (such as National Media Group, owned by Yuri Kovalchuk and his financial structures).

Such classification makes the task of separating media outlets into "oppositional" and "state-controlled" very difficult because formal ownership does not necessarily mean the degree of editorial independence.

In some works, Kiriya (2017) points out at least three main loyal oligarchs connected personally with the president (Yuri Kovalchuk who owns National media group, Alisher Usmanov who controls Mail.ru Group, and Kommersant and Yuri Berezkin who own RBC and *Komsomolskaya Pravda*). At the same time, it is important to say that in 2012, some loyal Kremlin oligarchs have been allowed to own critical media outlets, ready to present alternative critical points of view and covering activities of opposition leaders (notably Alexey Navalny, Serguey Udaltsov, Boris Nemtsov, and some other opposition leaders). Among such media, we can find some commercial news sites which are relatively independent such as the RBC Group, Echo Moskvy radio station (owned by Gazprom but editorially independent), and Kommersant, which has been recently acquired by Alisher Usmanov but was oriented towards being more provocative, addressing a more oppositional, well-educated public. The same can be said for both online media operated by well-known liberal RuNet activist Anton Nossik: Gazeta.ru and Lenta.ru. The entire table with classification is in the Supplementary File.

In Table 1, it can be observed that in 2012, the total average monthly reach of oppositional outlets on the internet represented around 50%. If we associate it with some state-controlled outlets that objectively covered the Moscow uprisings (Ria Novosti, for instance) we will obtain a much bigger figure.

In 2020, the situation changed drastically. The most important changes were structural and related to taking the most important internet outlets under financial or editorial control. Ria Novosti was restructured and became a part of the big propagandist holdings *Rossia Segodnya* [Russia Today] controlling pro-state networks of web-portals and radio stations outside Russia. The Kremlin forced the owner of Lenta.ru to change its editorial staff. Rbc.ru, known for its journalistic investigations and owned by oligarch Michail Prokhorov, changed the editorial team of Elizaveta Ossetinskaya under pressure from the Kremlin (Seddon, 2016) and later changed the owner to Grigory Beriozkin, the loyal oligarch who already controlled the big popular newspaper *Komsomolskaya Pravda* (kp.ru). In 2016, in response to foreign sanctions, the Russian Duma adopted a law limiting the foreign ownership of any media to 25%, which finally led to the great departure of foreign media owners from the print media market (Kiriya, 2017). It changed the ownership of nearly all important critical media outlets such as *Forbes* (previously owned by Axel Springer) and Vedomosti, where the new editor-in-chief, loyal to Russian state oil company Rosneft, provoked a change of editorial staff in 2020 (Seddon, 2020).

All such changes provoked great "alternativisation" of the public sphere, while the expelled editorial teams often created their own media outlets (e.g., former editor-in-chief of RBC, Elizaveta Ossetinskaya, created thebell.io, and Vedomosti staff opened vtimes.io). However, just one medium among them attained real success in terms of audience, Meduza.io. Thus, "alternativisation" means marginalisation and does not represent a considerable risk for the Kremlin.

In parallel to taking control of major media platforms, which might represent the alternative opinion, the state power considerably enlarged its presence in the internet space. Such big state online media as Rt.com, M24.ru, and Tass.ru started to acquire bigger audiences. All such outlets represent just web news versions of other known media. M24 is a subsidiary of Moscow 24 television station (controlled by state-owned VGTRK), Rt.com is under the control of Russia Today, and Tass.ru is the web version of the great state-owned information agency Tass.

In addition, we can see a rise in state-owned online media. In 2017, RBC published an investigation about the so-called "media factory," an informal group of reactionary online media sharing the same building and common investors with the legendary "troll factory" in Saint Petersburg (so-called "Agency of internet research" organizing troll propagandistic anti-opposition campaigns which is one of the Russian organizations accused by special prosecutor Robert Mueller in intervention into US elections). All such online media collected more than 30 million users in RuNet (Zakharov & Rousiaeva, 2017). After denying any connections with the "media factory" Evgeni Prigozhine, the oligarch close to Putin and owner of the "troll factory," transformed the "media factory" into media holding "patriot media," connecting such online media as Polit.info, Politpuzzle.com, and Riafan.ru (called Federal Agency of News). The editor-in-chief of the Federal Agency of News in his interview with Andrei Loshak described his work as "working in the context of the information defence" against the West (Loshak, 2020).

Table 1. Total monthly reach and share of opposition and state-controlled online media.

	2012		2020	
	Total monthly reach (thousands)	Share of total monthly reach %	Total monthly reach (thousands)	Share of total monthly reach %
Opposition	38,482.8	49	3,596	5
State owned	40,052	51	71,641	95

Notes: Echo.msk.ru and novayagazeta.ru have been excluded from the coverage of the database since 2016. Some figures on media outlets (such as Meduza) were unavailable for 2020—thus, the most recent data was used. Source: Built based on Web Index database provided by the official media measuring company Mediascope (before 2016, company TNS Russia; Mediascope, 2021).

The usage of original Web Index data for 2020 was not so relevant for comparison just because the Web Index is the database that includes media outlets based on their willingness to be measured. This means that some media outlets' coverage by measurements is not stable. Some media (and notably some relatively new media outlets including state-owned outlets) have never been covered by measurements. Therefore, we used the data of monthly visits provided by the media analytical tool Similar Web.

Such strategies allow state-controlled media to dominate in the online realm. In Table 2, we can see that such media accumulate more than one billion (1,020,698 thousand) visits monthly. It is emblematic that state-owned online media commonly not only use the alternative model of distribution, which is based on non-organic traffic, but they also make use of referral traffic, where users come to the pages by clicking links promoted on social media or search engines. A higher share of organic traffic means that users come to the web page of the media outlets by themselves because they trust such media and visit them consciously. A higher share of non-organic traffic means that users click on links from social media, search engine links, and aggregators. Such users occasionally come to the web page attracted by aggressive headlines. Such sites usually publish conspiracy theories, non-checked facts, and very dubious editorials. Thus, such media exploit the curiosity and occasional attention of mass users who are not very familiar with fact-checking and basic media literacy. Thus, we are calling such a strategy "littering the information space" with different kinds of propagandistic trash to increase the total traffic on state loyal internet media to make the pro-state discourse and topics largely dominate the internet.

To increase the presence of the state online news on the internet, the state adopted the so-called "Lugovoi law" (named after the deputy who proposed it), according to which news aggregators become responsible for the news they are aggregating on their top pages. Thus, search engines (starting with Yandex, the biggest) became responsible for the aggregated content coming from internet news that were not registered as mass media in Ruskomnadzor, the Russian internet watchdog. Eventually, it hugely transformed the key sources indexed by Yandex and almost eliminated alternative media from its top news (Daucé, 2017). Together with the strategy of "littering," the control over the Yandex algorithm gives the Russian state the ability to maximise attention on pro-state discourse.

We argue that all such strategies are oriented towards making state-manipulated and controlled news prevail in internet space, including social media. This corresponds to the third model of Deibert-Rohozinski (Deibert & Rohozinski, 2010), which is aimed at proposing competing content by the state.

5. Conclusion

In this article, all described methods of internet control in Russia have been put together to find a common logic between them. For a long time, the internet was interpreted as opposing state control, a liberal means of self-expression, and consequently a kind of parallel opposition discourse. Such a vision was inspired by the huge difference between an offline audience of traditional media and an online audience representing a more educated, critically thinking public. Through the present analysis, it has been demonstrated that such a vision does not correspond to the current digital mass reality, where most of the Russian population now has access to the internet. As a result, if 10 years ago the internet mainly attracted oppositional discourses, in the current situation it represents just a mini-model of the mainstream media, where the state has predominated since the mid-2000s. We may have called such processes the "alternativisation" of the internet space ten years ago, but today they represent its "mainstreamisation."

As a result, control over the internet in Russia has changed considerably over the last eight years since the Moscow 2011–2012 uprisings, and most structural measures are related to the massification of the internet and involvement in broader parts of the Russian media audiences, which makes mainstream media more visible inside internet space. This finally led to the structural measures oriented towards the "mainstreamisation" of the internet. Together, the balance between direct prohibition measures and structural measures ensures the Kremlin has control over the total media system, including the internet.

Kiriya (2014) formulated the main strategy of internet control in Russia as based on a gatekeeping function on the borders between different clusters of parallel and mainstream public spheres. As a result of the massification of internet media, such a strategy may be reconsidered. The borderline between mainstream and parallel public spheres is passing inside the internet. At the same time, the internet is not losing its status as a platform for "opposition" projects since the bargaining costs are low. As a result, oppositional political forces, as well as differ-

Table 2. Total monthly visits and average share of organic traffic for state-controlled and opposition outlets in 2020.

	Opposition online media	State-controlled online media
Total monthly visits (thousands)	100,107	1,020,698.75
Average share of organic (natural) traffic (%)	62.72	39.54

Note: Built-in Similar Web analytical tool based on March 2020 data from all main online outlets.

ent radical movements rejected by mainstream media, still use the internet to create their own media spaces. It makes the application of the term "alternative" media more difficult (in terms of self-organised or grassroots).

As we can see, the model of internet control in Russia combines direct blocking measures and the promotion of more structural measures oriented towards making state-produced online news prevail in the online information space. It is maintained by some ownership-related issues (such as the acquisition of a larger part of alternative media by loyal oligarchs), the development of state-owned information resources, and legal measures (such as influencing search engine news aggregation). Such measures were developed in parallel with some self-censorship measures: Making social media users afraid to comment in ways considered inappropriate by the state. Such methods represent a kind of balance between direct prohibition and self-control, which addresses our first hypothesis.

The utilised model of internet control and its evolution clearly distinguish Russian strategies of internet control from their more authoritarian analogues and notably from its analogues in the post-Soviet space. In this article, we made a clear distinction between Russian internet control and the Chinese model based on direct blocking and filtering measures. Thus, Russia differs considerably from other countries in post-Soviet spaces using the same measures—especially in the case of Turkmenistan and Uzbekistan. The findings of this article clearly show that the internet (and social media) should no longer be regarded as an oppositional or protest space, but as a part of the whole media landscape oriented towards maintaining the status quo. Here, we suggest a closer analysis of post-Soviet countries such as Kazakhstan or Belarus, which are much more like Russia in their models of control. In recent protests in Belarus, some analysts preferred to continue the "emancipating discourse of internet" (Bush, 2020). However, since summer 2020, the Belarussian regime does not seem to have demonstrated any willingness to change, so a more detailed analysis of Belarussian internet space and internet control is needed.

Acknowledgments

This article is an output of a research project implemented as part of the Basic Research Program at the National Research University Higher School of Economics (HSE University).

Conflict of Interests

The author declares no conflict of interest.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Bailey, O. G., Cammaert, B., & Carpentier, N. (2008). *Understanding alternative media*. McGraw Hill; Open University Press.

Becker, J. (2004). Lessons from Russia: A neo-authoritarian media system. *European Journal of Communication*, *19*(2), 139–163.

Bode, N., & Makarychev, A. (2013). The new social media in Russia. *Problems of Post-Communism*, *60*(2), 53–62.

Bodrunova, S. S., & Litvinenko, A. A. (2015). Four Russias in communication: Fragmentation of the Russian public sphere in the 2010s. In M. Glowacki & B. Dobek-Ostrowska (Eds.), *Democracy and media in Central and Eastern Europe 25 years on* (Vol. 4, pp. 63–79). Peter Lang.

Budnitsky, S., & Jia, L. (2018). Branding internet sovereignty: Digital media and the Chinese-Russian cyberalliance. *European Journal of Cultural Studies*, *21*(5), 594–613.

Bush, D. (2020, August 28). No modest voices: Social media and the protests in Belarus. *Stanford Freeman Spogli Institute for International Studies Blog*. https://fsi.stanford.edu/news/no-modest-voices-social-media-and-protests-belarus

Daucé, F. (2017). Political conflicts around the internet in Russia: The case of Yandex Novosti. *Laboratoriu:. Russian Review of Social Research*, *9*(2), 112–132.

Daucé, F., Loveluck, B., Ostromooukhova, B., & Zaytseva, A. (2019). From citizen investigators to cyber patrols: Volunteer internet regulation in Russia. *Laboratorium*. *Russian Review of Social Research*, *11*(3), 46–70.

Degtereva, E., & Kiriya, I. (2010). Russian TV market: Between state supervision, commercial logic, and simulacrum of public service. *Central European Journal of Communication*, *1*(4), 37–51.

Deibert, R., & Rohozinski, R. (2010). Control and subversion in Russian cyberspace. In R. Deibert, J. Palfrey, R. Rohozinski, & J. Zittrain (Eds.), *Access controlled: The shaping of power, rights, and rule in cyberspace* (pp. 15–34). MIT Press.

Deloitte. (2020). *Mediapotreblenie v Rossii 2020* [Media consumption in Russia 2020]. https://www2.deloitte.com/ru/ru/pages/technology-media-and-telecommunications/articles/media-consumption-in-russia.html#

Denisova, A. (2017). Democracy, protest, and public sphere in Russia after the 2011–2012 anti-government protests: Digital media at stake. *Media, Culture & Society*, *39*(7), 976–994.

Etling, B., Alexanyan, K., Kelly, J., Faris, R., Palfrey, J., & Gasser, U. (2010). *Public discourse in the Russian blogosphere: Mapping RuNet politics and mobilization public discourse in the Russian blogosphere*. Berkman Klein Center. https://cyber.harvard.edu/publications/2010/Public_Discourse_Russian_Blogosphere

Fuchs, C. (2010). Alternative media as critical media. *European Journal of Social Theory*, *13*(2), 173–192.

Fuchs, C. (2014). *Social media: A critical introduction*. SAGE.

Gabdulhakov, R. (2020). (Con)trolling the web: Social media user arrests, state-supported vigilantism, and citizen counter-forces in Russia. *Global Crime*, *21*(3/4), 283–305.

Goryashko, S., & Prokofiev, A. (2012, June 6). Hipster, kto ty? [Hipster, who are you?]. *Rossijskaya Gazeta*. https://rg.ru/2012/06/06/hipsters.html

Karatzogianni, A., Miazhevich, G., & Denisova, A. (2017). A comparative cyberconflict analysis of digital activism across post-Soviet countries. *Comparative Sociology*, *16*(1), 102–126.

Kiriya, I. (2012). The culture of subversion and Russian media landscape. *International Journal of Communication*, *6*(1), 446–466.

Kiriya, I. (2014). Social media as a tool of political isolation in the Russian public sphere. *Journal of Print and Media Technology Research*, *3*(2), 131–138.

Kiriya, I. (2017). The impact of international sanctions on Russia's media economy. *Russian Politics*, *2*(1), 80–97.

Kiriya, I. (2019). New and old institutions within the Russian media system. *Russian Journal of Communication*, *11*(1), 6–21.

Kiriya, I., & Sherstoboeva, E. (2015). Russian media piracy in the context of censoring practices. *International Journal of Communication*, *9*(1), 839–851.

Koltsova, O. (2006). *News media and power in Russia*. Routledge.

Koltsova, O., & Shcherbak, A. (2015). "LiveJournal Libra!": The political blogosphere and voting preferences in Russia in 2011–2012. *New Media & Society*, *17*(10), 1715–1732.

Loshak, A. (2020). InterNYET: A history of the Russian internet—Episode #5. [Video]. https://www.youtube.com/watch?v=9Ep4tG7fapg

McCombie, S., Uhlmann, A. J., & Morrison, S. (2020). The US 2016 presidential election & Russia's troll farms. *Intelligence and National Security*, *35*(1), 95–114.

Mediascop. (2021). *Auditoria Interneta v Rossii v 2020 godu* [Internet audience in Russia in 2020]. https://mediascope.net/news/1250827

Mickiewicz, E. (2008). *Television, power, and the public in Russia*. Cambridge University Press.

Oates, S. (2007). The neo-Soviet model of the media. *Europe-Asia Studies*, *59*(8), 1279–1297.

Oates, S. (2016). Russian media in the digital age: Propaganda rewired. *Russian Politics*, *1*(4), 398–417.

Public Opinion Foundation. (2011). *Internet v Rossii* [Internet in Russia]. https://fom.ru/uploads/files/Internet_Russia_Summer_2011.pdf

Remmer, V. (2017). The role of internet based social networks in Russian protest movement mobilization. *Central European Journal of International and Security Studies*, *11*(1), 104–135.

Roth, A. (2019, April 28). Russia's great firewall: Is it meant to keep information in—Or out? *The Guardian*. https://www.theguardian.com/technology/2019/apr/28/russia-great-firewall-sovereign-internet-bill-keeping-information-in-or-out

Russian Federation. (2019a). *Federal'nyj zakon "O vnesenii izmenenij v Federal'nyj zakon "O svyazi" i Federal'nyj zakon "Ob informacii, informacionnyh tekhnologiyah i o zashchite informacii" ot 01.05.2019 N 90-FZ* [Federal law on changes in the federal law "on communication" and the federal law "on information, information technologies and information protection" from 01.05.2019 N 90-FZ]. http://www.consultant.ru/document/cons_doc_LAW_323815

Russian Federation. (2019b). *Federal'nyj zakon "O vnesenii izmenenij v stat'yu 15.3 Federal'nogo zakona "Ob informacii, informacionnyh tekhnologiyah i o zashchite informacii" ot 18.03.2019 N 31-FZ* [Federal law on changes in article 15.3 of the federal law "on information, information technologies and information protection" from 18.03.2019 N 31-FZ]. http://www.consultant.ru/document/cons_doc_LAW_320401

Russian Federation. (2020). *Federal'nyj zakon "O vnesenii izmenenij v Federal'nyj zakon "Ob informacii, informacionnyh tekhnologiyah i o zashchite informacii" ot 30.12.2020 N 530-FZ* [Federal law on changes in the federal law "on information, information technologies and information protection" from 30.12.2020 N 530-FZ]. http://www.consultant.ru/document/cons_doc_LAW_372700/3d0cac60971a511280cbba229d9b6329c07731f7

Russian Federation. (2021). *Ugolovny Kodeks Rossijskoj Federatsii ot 13.06.1996 N 63-FZ (s izm. 05.04.2021)* [Penal code of Russian Federation from 13.06.1996 N 63 FZ (with changes from 05.04.2021)]. http://www.consultant.ru/document/cons_doc_LAW_10699

Schimpfossl, E., & Yablokov, I. (2014). Coercion or conformism? Censorship and self-censorship among Russian media personalities and reporters in the 2010s. *Demokratizatsiya*, *22*(2), 295–312.

Seddon, M. (2016, May 18). Editors at Russia's RBC media group sacked after Putin article. *Financial Times*. https://www.ft.com/content/45a8a9c4-1c3b-11e6-b286-cddde55ca122

Seddon, M. (2020, May 31). Pro-Kremlin entrepreneur buys leading Russian business newspaper. *Financial Times*. https://www.ft.com/content/ce7114dd-cbcc-4292-8a5a-199fe3688a96

Surganova, E. (2014, September 15). Galina Timchenko: Nikto iz nas ne mechtaet delat "Kolokol" [Nobody from us dreams to produce "The Bell"]. *Forbes*. https://www.forbes.ru/kompanii/internet-telekom-i-media/267611-galina-timchenko-nikto-iz-nas-ne-mechtaet-delat-kolokol

Toepfl, F. (2011). Managing public outrage: Power, scandal, and new media in contemporary Russia. *New Media & Society*, *13*(8), 1301–1319.

Toepfl, F. (2018). From connective to collective action: Internet elections as a digital tool to centralize and formalize protest in Russia. *Information Communication and Society*, *21*(4), 531–547.

Vartanova, E. (2011). The Russian media model in the context of post-Soviet dynamics. In D. C. Hallin & P. Mancini (Eds.), *Comparing media systems beyond the western world* (pp. 119–142). Cambridge University Press.

Vendil Pallin, C. (2017). Internet control through ownership: The case of Russia. *Post-Soviet Affairs*, *33*(1), 16–33.

Zakharov, A., & Rousiaeva, P. (2017, March 24). Rassledovanye RBK: Kak iz "fabriki trollej" vyrosla "fabrika media" [RBC's investigation: How the "trolls factory" transformed into "media factory"]. *RBK journal*. https://www.rbc.ru/magazine/2017/04/58d106b09a794710fa8934ac

Zassoursky, I. (2016). *Media and power in post-Soviet Russia*. Routledge.

About the Author

Ilya Kiriya graduated in journalism from Moscow State University and Grenoble University (Master of Research). He holds a PhD in journalism (2002) from Moscow State University and a PhD in information and communication (2007) from Grenoble Stendhal University (France). He is professor and head of the Media School at HSE University in Moscow. His main interests focus on the political economy of media and communication in post-Soviet Russia and the post-Soviet public sphere.

Media and Communication (ISSN: 2183–2439)
2021, Volume 9, Issue 4, Pages 27–38
https://doi.org/10.17645/mac.v9i4.4237

Article

The Vulnerabilities of Trusted Notifier-Models in Russia: The Case of Netoscope

Liudmila Sivetc [1,*] and Mariëlle Wijermars [2]

[1] Faculty of Law, University of Turku, Finland; E-Mail: liusiv@utu.fi
[2] Faculty of Arts and Social Sciences, Maastricht University, The Netherlands; E-Mail: m.wijermars@maastrichtuniversity.nl

* Corresponding author

Submitted: 28 February 2021 | Accepted: 25 June 2021 | Published: 21 October 2021

Abstract
Current digital ecosystems are shaped by platformisation, algorithmic recommender systems, and news personalisation. These (algorithmic) infrastructures influence online news dissemination and therefore necessitate a reconceptualisation of how online media control is or may be exercised in states with restricted media freedom. Indeed, the degree of media plurality and journalistic independence becomes irrelevant when reporting is available but difficult to access; for example, if the websites of media outlets are not indexed or recommended by the search engines, news aggregators, or social media platforms that function as algorithmic gatekeepers. Research approaches to media control need to be broadened because authoritarian governments are increasingly adopting policies that govern the internet *through* its infrastructure; the power they leverage against private infrastructure owners yields more effective—and less easily perceptible—control over online content dissemination. Zooming in on the use of trusted notifier-models to counter online harms in Russia, we examine the Netoscope project (a database of Russian domain names suspected of malware, botnet, or phishing activities) in which federal censor Roskomnadzor cooperates with, e.g., Yandex (that downranks listed domains in search results), Kaspersky, and foreign partners. Based on publicly available reports, media coverage, and semi-structured interviews, the article analyses the degree of influence, control, and oversight of Netoscope's participating partners over the database and its applications. We argue that, in the absence of effective legal safeguards and transparency requirements, the politicised nature of internet infrastructure makes the trusted notifier-model vulnerable to abuse in authoritarian states.

Keywords
authoritarian states; internet governance; internet sovereignty; news personalisation; Netoscope project; platformisation; Roskomnadzor; Russia; trusted notifier-model

Issue
This article is part of the issue "Media Control Revisited: Challenges, Bottom-Up Resistance and Agency in the Digital Age" edited by Olga Dovbysh (University of Helsinki, Finland) and Esther Somfalvy (Research Centre for East European Studies at the University of Bremen, Germany).

1. Introduction

Current digital ecosystems are shaped by platformisation, algorithmic recommender systems, and—increasingly—news personalisation (van Dijck, 2020). These (algorithmic) infrastructures influence the online dissemination of news and therefore necessitate a reconceptualisation of how online media control is or can be exercised in states with restricted media freedom. Indeed, the degree of media plurality and journalistic independence becomes irrelevant when reporting is available but difficult to access; for example, if the websites of media outlets are not indexed or recommended by the search engines, news aggregators, or social media platforms that function as algorithmic gatekeepers (Napoli, 2015). This is all the more important since authoritarian governments increasingly adopt policies that govern the internet *through* its infrastructure

(Sivetc, 2021) and use the power of private infrastructure owners to achieve effective, but less easily perceptible, control over online content dissemination. For example, research indicates that, in Russia, Yandex's search engine and news aggregator demonstrate a bias and "referred users to significantly fewer websites that contained information" about protests (Kravets & Toepfl, 2021, p. 1). Zooming in on a concrete case where a trusted notifier-model (Schwemer, 2019) has been employed to counter online harm in Russia, we argue that, in absence of effective legal safeguards, accountability mechanisms and transparency requirements, the politicised nature of internet infrastructures makes this model vulnerable to abuse in authoritarian states.

Governments increasingly seek to control their "national" digital spaces by introducing online content regulations and expanding their influence over critical internet resources, such as Domain Name Systems (DNS; Mueller, 2010). When critical internet resources belong to private companies or (non-profit) organisations, governments, therefore, seek to cooperate with or co-opt them to decide inter alia on the accessibility of online content (Balkin, 2014). For instance, establishing control over the national DNS infrastructure enables one to control connectivity among internet users: The DNS system, similar to a telephone book, connects names (URLs) with corresponding numbers (IP address where the resource is hosted) and therefore serves as "a necessary prelude to communication" (Klein, 2002, p. 195). Since country-code top-level domains (ccTLDs) such as .ru, are governed by relevant authorities at the national level, studying the relations between national governments, private parties, and not-for-profit organisations in this sphere is important as they determine the availability of online content (Schwemer, 2018). Moreover, national domain name registries, government bodies, and various private or public partners can be involved in online content control by creating "trusted notifier-models" for flagging suspicious domain names (Schwemer, 2019, p. 3).

Russia, a country in which media freedom is significantly restricted, actively seeks to expand its control over internet infrastructure and thereby strengthen its capacity to censor online content (Sivetc, 2021; Wijermars, 2021). Under the 2019 Russian Internet Sovereignty Act, Russia became one of the stakeholders of its national registry (the Coordination Center for top-level domains .ru and .рф) in June 2020 (Coordination Center, 2020). Responsible for, among other tasks, the allocation and deallocation of domain names, the Coordination Center occupies a powerful position which may be a valuable asset in its cooperation with other stakeholders, including the Russian state. This article analyses to what extent a governance model which relies on trusted notifiers, and in which the Russian internet regulator Roskomnadzor (Federal Service for Supervision of Communications, Information Technology, and Mass Media) cooperates with the Coordination Center and other key internet infrastructure owners, could be used for alternate ends (by the regulator or other actors). Our argument builds upon an examination of the Netoscope project (a database of .ru domain names suspected of malware, botnet, or phishing activities). The project was launched by the Coordination Center in 2012 and, by 2021, involved 17 partners, including Roskomnadzor, who contribute to Netoscope's database of harmful domain names, thereby affecting their reputation and, potentially, their algorithmic ranking.

Previous studies of Russian internet governance concerning media control have focused on federal legislation, media ownership structures, censorship, and surveillance (Litvinenko & Toepfl, 2019; Lokot, 2018; Sherstoboeva, 2020; Vendil Pallin, 2017). The "infrastructural turn" in internet governance scholarship (Musiani et al., 2016) has only recently started to be addressed with regard to Russia (Daucé & Musiani, 2021). Control mechanisms that function through infrastructures and the governance models involved have yet to be substantially investigated. This is of particular relevance as Russia seeks to create a "sovereign" internet whose successful realisation relies on state control over the Russian internet's infrastructure, including the creation of a national DNS (Stadnik, 2021).

To address this gap and demonstrate the need to complement existing approaches to studying media freedom with research into the governance of the algorithmic and physical infrastructures that shape online news dissemination, we examine the relations between the Coordination Center and the various partners it collaborates with within Netoscope. Based on publicly available reports, media coverage, and semi-structured interviews with representatives from the Coordination Center, Kaspersky (national partner), and SURFnet (international partner), we seek to understand the nature and dynamics of the trusted notifier-model which underlies the partnership and to explore the extent to which various Netoscope partners can influence, control, and have insight into the database and its applications. We interpret the implications of the governance structures we uncover and argue that, as a result of limited transparency, this governance model may be vulnerable to manipulation or abuse towards media control or other restrictive objectives.

2. The Place of Trusted Notifier Systems in Internet Governance

The introduction of state regulation of online content is by now a common trend across political systems, illustrating a gradual shift in the balance away from the multistakeholder approach, long held to be inherent to internet governance, towards more state-centred tendencies. The multistakeholder approach emphasises the global nature and complex interdependence of the internet; its governance therefore should involve not only states but also businesses, civil society, and communities of technical experts (Dutton, 2015). In contrast, the

state-centred approach to internet governance, referred to as internet balkanisation (Hill, 2012), fragmentation (Drake et al., 2016), or sovereignisation (Möllers, 2021) focuses on state regulation or "self-determination" with regard to local internet arenas. Notwithstanding the push towards sovereignisation, self-regulatory models continue to be prevalent, for example in efforts to limit the dissemination of illegal and harmful content on social media platforms.

In this context, trusted notifier-models have emerged as a way to disable access to illegal online content on the basis of notices sent by "trusted flaggers" or "trusted notifiers" (Schwemer, 2019, p. 2). This expertise can come from individuals, private organisations, civil society organisations, semi-public bodies, and public authorities (Schwemer, 2019, p. 3). For example, the trusted notifier-model is supported by the European Commission as it encourages platforms to collaborate with public authorities and trusted notifiers to take down illegal content (European Commission, 2017). Although trusted notifiers can act in different contexts (from flagging terrorist speech to identifying copyright infringements) several general features of such governance models can be identified (Schwemer, 2019): (1) Trusted-notifier models emerge as voluntary arrangements; (2) trusted notifiers act as privileged parties with a direct channel to the intermediary that has the capacity to affect the accessibility of flagged content; (3) there is no requirement of preliminary judiciary assessment of content flagged by trusted notifiers; and (4) as a form of privatised enforcement, the model suffers from a democratic deficit and can be challenged from the perspective of the rule of law, legal certainty, accountability, right to due process, as well as freedom of expression. In the context of initiatives aimed at countering disinformation, for example, outsourcing decisions on politically contentious issues to trusted notifiers may result in overcensoring with limited or no opportunity for redress.

Various international examples exist of the creation of public-private partnerships with the specific aim of countering malware and botnets, similar to the case under examination in this article (Dupont, 2017). Examining such anti-botnet initiatives launched between 2005 and 2010 in Australia, Japan, South Korea, Germany, and the Netherlands, Dupont (2017) explains that they centre around the engagement of internet service providers (ISPs) and anti-virus companies, typically encompassing private entities who are each other's direct competitors and are "often implemented by public Internet regulatory agencies attached to economic development and telecommunications ministries" (Dupont, 2017, p. 109). At their core is the establishment of information-sharing systems between telecommunications regulatory agencies and ISPs to aggregate data on botnets and identify infected devices. In South Korea, the Netherlands, and the United States, ISPs are known to place infected machines, whose users are "unable or unwilling to rectify the situation" in a "digital quarantine" by disrupting their internet access until the infection has been addressed (Dupont, 2017, p. 109); as a form of private enforcement, such practices give rise to legal and ethical concerns.

3. Russian Internet Governance and Media Control

Up until 2012, the Russian state demonstrated a relatively hands-off approach regarding internet regulation. Rather than employing filtering, restricting internet access, or blocking online content, the online domain was governed through more subtle means as Russia sought to shape online discourses "through effective counterinformation campaigns that overwhelm, discredit, or demoralize opponents" (Deibert & Rohozinski, 2010, p. 27). Therefore, the internet was able to function as a counterweight to the increasingly restricted traditional media (federal television, newspapers) and flourish as a platform for independent journalism and political activism (Wijermars & Lehtisaari, 2020). Russia had already taken several "preparatory steps" by enhancing state ownership of internet companies, attaching the status of mass media (and thereby the restrictions applicable to them) to their online counterparts, and floating the first proposals to establish a "national firewall" (Lonkila et al., 2020).

Since 2012, Russia intensified internet control, for example, by introducing website blocking legislation. Roskomnadzor was established in 2008 to regulate mass media and telecommunications and issue licences, and has since played a central role in website blocking procedures (Sivetc, 2020). In June 2020, the European Court of Human Rights criticised this practice when it ruled in two separate cases (*Kharitonov v. Russia*, 2020; *OOO Flavus and others v. Russia*, 2020) that Russia's website blocking legislation violates Article 10 of the European Convention on Human Rights. The Court found that the legal framework for website blocking jeopardises freedom of expression. It grants Roskomnadzor the ability to, without preliminary court oversight, block access not only to the allegedly unlawful content but also the entire website on which any such content is published (in these cases, e.g., grani.ru, an oppositional online media outlet). Moreover, implementation procedures affect innocent websites hosted on the same server as the targeted website (on overblocking and the ban of messenger Telegram, see Ermoshina & Musiani, 2021). Roskomnadzor's prerogatives in restricting access to online content without preliminary court oversight are expanding. The federal agency also partakes in extralegal internet governance practices, as is the case in the example we examine.

The technical obstacles Roskomnadzor encountered in putting in place effective website blocking (Ermoshina & Musiani, 2021; Stadnik, 2021) have led to the restructuring of Russian internet governance through the Russian Internet Sovereignty Act (2019). This law transferred the implementation of website blocking from

ISPs to the state. Through the obligatory placement of devices equipped with deep packet inspection technologies, Roskomnadzor was empowered to directly and more accurately filter and block websites, which should limit overblocking. However, lessening the dependence on ISPs and using state-controlled deep packet inspection filters may turn the website blocking mechanisms into a black box that is non-transparent to public and providers' scrutiny (Stadnik, 2021).

In addition to controlling online speech through legislative measures, the Russian government has co-opted internet gatekeepers to use their private rules to affect online content (Daucé & Loveluck, 2021). Their efforts to control which news items and sources are recommended by news aggregators, resulting in the law "On News Aggregators" (Wijermars, 2021), clearly indicate that the authorities are aware of the centrality of platforms and algorithmic infrastructures in online news dissemination. Empirical research suggests that Yandex's search engine and news aggregator indeed "forwar[d] users to fewer websites that regularly featured criticism of Russia's authoritarian leadership" (Kravets & Toepfl, 2021, p. 1). Yet, within scholarship on media freedom in Russia, the role of these intermediaries and governmental efforts to control them has received limited scrutiny. While for many Russian technological companies, their degree of independence vis-a-vis the Russian state has been (rightfully) questioned, the emergence of trusted notifier-models (as exemplified by Netoscope) within Russian internet governance and its possible implications necessitates further scrutiny as both part of and separate from the general sovereignisation trend.

4. Methodology

To gain insight into Netoscope and its governance structure we triangulated multiple sources. First, we analysed Coordination Center reports (2013–2020) that contain a section dedicated to counteracting illegal activities that use domains .ru/.рф, providing concise, general information about Netoscope and its main achievements. Second, we examined media coverage using the INTEGRUM Profi database, which provided additional information on the development of Netoscope, its partners, and the applications of the database. We queried the database with the Russian project name (*Нетоскоп*) for the period 1 January 2011–30 September 2020. Upon manually assessing relevance and removing duplicates, this resulted in 48 unique results. Media coverage was most frequent in 2013 (11 unique results) when the project's first results were published, and 2018 (10 results) in connection to the project's collaboration with FIFA. A substantial number (16) concerned publications by IT websites and magazines. Overall, media coverage can be characterised as being limited in frequency and largely guided by press releases.

Third, we conducted semi-structured interviews with Netoscope partners; all partners were invited, yet only three accepted the invitation. We interviewed a representative from the Coordination Center, who requested anonymity; Andrey Yarnykh, the Director of the Strategic Development Project of Kaspersky in Russia; and Roland van Rijswijk-Deij, who was employed as a researcher at SURFnet at the time when their agreement with Netoscope was signed and involved in the coordination of the collaboration. Each interviewee was asked to answer the same set of pre-prepared questions. All interviews were conducted online, in January and August 2020. This interview guide included several groups of questions: general questions regarding Netoscope; questions related to the motivations for joining and the role of the interviewee's organisation or company in Netoscope; questions about the relations among the project partners; questions about the Netoscope database (e.g., whether the interviewee's organisation contributes to the Netoscope database, uses it, has access to and control over it, ability to see which partner has flagged a certain domain name, whether it has any verification or safeguard mechanisms to prevent or remedy mistakes); and finally, a question concerning Roskomnadzor's participation in Netoscope. At the end of the interview, interviewees were invited to add anything else they would like to share regarding Netoscope. Since SURFnet's involvement in Netoscope is limited, this interview generated much less information and correspondingly features less prominently in our analysis. All translations were carried out by the authors.

The fact that many project partners declined our interview request presents a clear limitation to our study; yet, this is a condition that is commonly shared by research in this area focusing on Russia. For example, both Yandex, Russia's leading technology company, and Roskomnadzor are notoriously closed to information requests from researchers. Since the conducted interviews present three distinct perspectives (the Coordination Center, a Russian partner, and an international partner) we are nonetheless able to present a sufficiently comprehensive picture of how Netoscope functions. In interpreting these interviews, one also has to consider that, given the politicisation of internet infrastructure in Russia, interviewees may present an incomplete/one-sided view of the situation. Therefore, we compared and complemented findings with data from Coordination Center reports and media coverage whenever possible (again, taking into consideration the limitations in the availability and reliability of the latter sources). In the next section, we first present the insights gathered from these sources to tell a coherent story about the development and functioning of Netoscope. Since the way in which interviewees narrativise their positions is an important source for understanding the project, these statements are presented comprehensively. A critical discussion of the picture emerging from our sources then follows in Section 5.3.

5. Netoscope

5.1. History and Functionality of Netoscope

Netoscope was launched in 2012 by the Coordination Center. As stated on its official website, the project "aims at making the Russian domain space safer for users" (Coordination Center, 2012). In our interview, the representative of the Coordination Center, who is directly involved in the functioning of Netoscope, explained that the project was not intended for the regulation of the Russian internet; rather, it was deemed necessary for improving the reputation of the Russian top-level domains, since they did not rank among the safest domains in 2009–2011. Although this low ranking, according to the representative, lacked a proper justification, they admitted the validity of some of the security concerns; the .ru domain was indeed used for malicious activities, such as malware and the creation and operation of botnets. In the representative's view, these malicious activities may be explained by the low prices of domain name registration and the (according to them, incorrect) impression that the Coordination Center was indifferent to activities in the Russian ccTLDs. On the contrary, the representative emphasised that the Coordination Center was very much interested in making the domain safe for internet users but the issue was that the Coordination Center lacked the necessary competencies to identify domain names involved in malicious activities. Therefore, it proposed Netoscope as a platform for cooperating with cybersecurity experts.

Cybersecurity experts, in turn, needed the cooperation with the Coordination Center because only they are able to terminate the domain name delegation of resources involved in the "epidemic" dissemination of, for example, malware, as Andrey Yarnykh, the Director of Strategic Development Project of Kaspersky in Russia, indicated in the interview. The termination of the domain delegation does not cancel the registration of a domain name; it terminates the connectivity between the domain name and the corresponding address, which makes the respective website inaccessible until the delegation is restored. Experts employed by Kaspersky, he indicated, can detect malware being spread by such resources and identify which domain names are used for coordinating command points. To prevent such epidemics from developing, the resources behind them should be disabled directly by terminating the delegation of the domain names involved. Therefore, Kaspersky had an interest in being able to inform the Coordination Center on domain names engaged in malicious activities and request the termination of their delegation. According to Yarnykh, Netoscope provided the necessary mechanisms for that purpose and Kaspersky sends information on malicious domain names to the project to enable the Coordination Center to expeditiously react to cyberthreats. Here, his account differs from that provided by the Coordination Center representative, who pointed out that Netoscope is only the basis for technical and scientific collaboration. The termination of domain name delegation, which indeed lies within the mandate of the Coordination Center, is realised through a separate trusted-notifier mechanism that is more formalised and transparent in its procedure and in which Kaspersky and other Russian Netoscope partners are authorised to request undelegation.

According to Yarnykh, Netoscope effectively combats the viral spread of malware, botnets, and phishing by disabling coordinating command points, which decreases the levels of malicious activities in the Russian ccTLDs as well as globally. The Coordination Center representative also indicated that the cooperation within Netoscope has led to a decrease in the number of malicious activities in the Russian ccTLDs and thereby improved their reputation. If, in the beginning, Netoscope flagged a hundred thousand malicious domains per year, by 2020, the numbers had decreased significantly and the domain became "cleaner" (measuring the impact of such partnerships is, however, difficult; Dupont, 2017).

Yarnykh highlighted that Kaspersky does not gain commercial benefit from participating in Netoscope but acts as a "donor." The company's interest, he said, consists solely in contributing to stable internet development. To this aim, the company cooperates with Netoscope partners to make the Russian ccTLDs "cleaner and more protected." Kaspersky cooperates with partners involved in Netoscope outside of the project as well, but these processes are conducted "in different formats" than those within Netoscope.

The Coordination Center representative explained that Netoscope was created upon several meetings with experts. Some of them had shown interest in cooperating, others were specially invited by the Coordination Center. Initially, the representative indicated, Netoscope involved such partners as RU-CERT, Kaspersky, Group IB, and the Technical Center "Internet"; i.e., Russian cybersecurity companies. The Coordination Center's 2012 report indicates that Yandex, which can be considered the Russian counterpart and competitor of Google, providing a broad array of digital services, including internet browser, search engine and news aggregation, joined in 2012 (Coordination Center, 2013, p. 11). Gradually, additional companies and organisations also joined the project, including three foreign partners: IThreat Cyber Group (United States), SURFnet (the Netherlands), and FIFA.

Table 1 presents an overview of the 17 project partners listed on the website and indicates their main areas of activity. It shows that Netoscope differs from the anti-botnet public-private partnerships described by Dupont (2017) in several respects. First, while cybersecurity companies make up a substantive proportion of partners, the central role of ISPs Dupont identified is lacking. The only partner involved in providing internet services is Rostelecom. However, its membership may be explained by its involvement in the creation of the

Table 1. Netoscope partners.

Partner	Organisation type	Role within Netoscope
Coordination Center	Russian domain name registry organisation	Coordinator
Technical Center "Internet" (TCI)	Russian organisation maintaining the main registry for ccTLDs .ru, .рф and .su	Participant
Roskomnadzor	Russian federal executive authority for media and internet regulation	Participant
National Computer Incident Response and Coordination Center	Russian Computer Emergency Response Team responsible for the protection of governmental networks of the Russian Federation	Participant
RU-CERT	Russian autonomous non-profit organisation. Computer Emergency Response Team	Participant
Group IB	Russian private cybersecurity company	Participant
Kaspersky	Russian private cybersecurity company	Participant
SkyDNS	Russian private cybersecurity company	Participant
Dr. Web	Russian private cybersecurity company	Participant
BI-ZONE	Russian private cybersecurity company. Daughter company of Sber (previously, Sberbank)	Participant
MasterCard Members' Association	Russian non-profit organisation	Participant
Rostelecom	Russian private telecommunications company. Market leader in provision of (mobile) internet services	Participant
Yandex	Russian multinational corporation offering a wide array of digital services. Owner of Yandex browser and search engine	Participant
Mail.ru Group	Russian corporation active in email services, e-commerce, B2B, media, instant messaging. Owner of VKontakte and Odnoklassniki	Participant
IThreat	American private cybersecurity company	Participant
SURFnet	Cooperative association of Dutch educational and research institutions aimed at the development and procurement of information and communication technology facilities and knowledge sharing	Participant
FIFA	French non-profit organisation. Organiser of the FIFA World Cup	Participant

Russian browser and search engine Sputnik, launched in 2014, which filtered various harmful materials from its search results through its collaboration with Kaspersky, Netoscope, and Roskomnadzor ("'Sputnik' iskliuchaet iz poiskovoi," 2014). Because of its limited success, the Sputnik search engine was discontinued in 2020, yet the company continues to provide search solutions to corporate and government clients ("Poiskovik 'Sputnik' prekratil," 2020). Second, it includes two key players of the Russian internet: Yandex, previously introduced, and Mail.ru Group, which (among many other activities) is the owner of the popular social media platforms Vkontakte and Odnoklassniki and a (much less popular) search engine and news aggregator. Rambler Media Group, another prominent digital media company owned by Sber (a state-owned bank), is not included. Finally, there are three non-Russian partners, whose partnership appears to be motivated differently, as will be discussed below.

According to the Coordination Center's 2016 report, Roskomnadzor joined Netoscope on 19 April 2016 (Coordination Center, 2017, p. 12). Roskomnadzor and Netoscope concluded an agreement on cooperation aimed inter alia at "the joint investigation of content, types, and features of unlawful online information and the development of means of precluding it from dissemination on the Internet" (Coordination Center, 2016, p. 2), a formulation which suggests a scope that extends beyond botnets and malware. Despite becoming an official partner in 2016, Roskomnadzor, as the Coordination Center representative clarified, was involved in Netoscope from the very beginning and was an active participant both before and after concluding the agreement. Their cooperation practices were not affected by the changed status, the representative stated: "[T]here have not been any cardinal changes. Instead, there has been active, annual, everyday, on-time performance." Andrey Yarnykh also indicated that the practices

of cooperation within Netoscope did not change when Roskomnadzor joined; at least, Kaspersky did not notice any changes. The company continues to send information to the database in accordance with its own expertise: phishing, spam, and malware. Yarnykh assumes that Roskomnadzor, just as other partners, contributes to the project within the agency's expertise, in a way that benefits Netoscope's overall objective.

According to the Coordination Center representative, experts contribute to Netoscope by submitting information on domain names involved in phishing, malware, and botnet activities to a database that accumulates the information and stores all suspicious domain names. This means that once a domain name is included in the Netoscope database, it will never be excluded from it, even when the flagged domain name no longer hosts the malicious content. If the domain name ceases to exist (if its registration in one of the Russian ccTLDs is discontinued) this also does not affect the information stored in the database. These structural characteristics leave the issue of how to interpret the information about a domain's entry into the database up to the user of the database. The principle of permanent storage, the Coordination Center representative explained, is based on the assumption that a domain name that has been used for malicious activities in the past is likely to be used again and therefore retains its dangerous potential. Yet, it means there is no possibility for domains that have been falsely flagged or flagged as a result of manipulation (e.g., a malicious actor simulating an attack and connecting it to the domain of an opposition-related website) to rid themselves of the reputational damage and its possible consequences. The available information also suggests domain name owners are not necessarily informed if they are added to the database.

The Netoscope database serves as the basis for the Domain Checker available on the Netoscope website. Any internet user can use it to find out whether a domain name registered in the .ru, .su, or .рф domains has been flagged by Netoscope. For example, (oppositional) online media outlet grani.ru was blocked in March 2014 on the allegation of publishing calls to participate in unauthorised mass protests. The Domain Checker (Netoscope, 2021) indicates the following result for the domain: "On the domain name grani.ru project partners recorded the following malicious activities: Formerly Malware."

In December 2020, the Netoscope database contained approximately 4,7 million domain names (Netoscope, 2020). The Coordination Center representative explained that this figure should not be understood as an indicator of a high level of malicious activity. Only a small number of these, around five thousand, represent domain names flagged as "currently malicious." In the case of grani.ru, its inclusion in the database indicates that the domain name is outside the scope of currently malicious websites, yet possessed this status at some point in the past. This status should signal to users that the website is safe to access. However, the fact that it was previously flagged by Netoscope may also give rise to questions regarding the website's safety. For example, according to the representative, companies involved in the domain name business adjust their decision to purchase a certain domain name if it has been flagged by Netoscope.

Netoscope has another direct and intended effect: The Coordination Center's 2014 report states that Yandex, the provider of Russia's most popular search engine, has been using the Netoscope database since 2014 to exclude links to malicious websites from its search results (Coordination Center, 2015, p. 11; see also Kudriavtseva, 2020). The Coordination Center representative confirmed that Yandex can use the Netoscope database to adjust how its algorithms decide on which websites are prioritised in search results, yet stressed it is but one of many resources Yandex uses as an input source for its algorithms. Yandex also contributes to the database: According to the representative, the Yandex Safe Browsing database has been used by Netoscope to enrich and refine data about domain names included in the Netoscope database.

5.2. Netoscope as a Trusted Notifier-Model

The Coordination Center representative highlighted an important feature of Netoscope: The project facilitates collaboration among competitors. Most of the partners involved in Netoscope are commercial entities active in adjacent fields; therefore, they prefer not to share information with other (cybersecurity, technology) companies. Yet, as partners in Netoscope, they are willing to share information with the Coordination Center and contribute to the database. According to the representative, the partners cooperate because they share the common goal of making the Russian ccTLDs safer. Moreover, by cooperating, they develop "mechanisms" for identifying malicious activities, which enhances their competencies and thereby their competitiveness in the market. However, each Netoscope partner is unaware of what information the other partners share with Netoscope. As the Coordination Center representative explained, Group IB, for instance, does not know which domain names have been flagged as malicious by Kaspersky: The partners have agreed on this practice because, as competitors, they "do not support the idea that some of them donate the information, while the others only use it without contributing."

Kaspersky's Andrey Yarnykh also mentioned market competition among Netoscope partners as the reason for the fact that there is only unilateral communication between Netoscope and the company. He said that, because Kaspersky's databases with information on malware, phishing, and botnets are used in conducting its projects, this information should be kept secret from competitors. Although Kaspersky sends the information to Netoscope, this information is available only to the project but not to its partners. According to Yarnykh,

"it would be incorrect and wrong if Netoscope presented a resource that shares the information we provided." Rather, "Netoscope was initially designed as a resource to which the partners contribute information but do not take from it"; he also indicates that Netoscope accumulates information but does not disseminate it.

Another aspect affecting information-sharing practices within Netoscope is the different competencies respective Netoscope partners have, which, according to the Coordination Center representative, is noted in the agreement on cooperation. They explained the actual cooperation occurs as follows: The Netoscope database is located at the Coordination Center. Each partner submits information on those domain names that it identifies as being involved in malicious activities to the database. The representative stressed that partners decide whether to flag a domain name, in accordance with their particular expertise. Yarnykh indicated that Netoscope aggregates information sent by the partners and issues reports on the levels of malicious activity. These reports are purposely designed not to reveal the size and content of each partner's contribution to the project. As Yarnykh said, reports provide "statistics rather than analytics." Netoscope does not enable Kaspersky to see which partner flagged a certain domain name.

Importantly, as the Coordination Center representative indicated, Netoscope relies on partners' expertise and does not verify inputs into the database. They explained that such verification falls outside of the Coordination Center's remit and they do not employ experts to perform such verification checks. If a Netoscope partner "says that this domain name is connected with phishing at this moment, it means that the partner answers for [the accuracy of] its words."

According to the Coordination Center representative, Netoscope also relies on the partners' expertise in deciding on notifications about malicious activities received from internet users. Users can inform Netoscope by pressing the button "report malware" on the Netoscope website. Netoscope then sorts out notifications about botnets, phishing, and malware and forwards this information to the relevant partner specialising in identifying the respective malicious activity. Netoscope has received many complaints on malicious activities from users, the representative mentioned, without specifying whether any NGOs or organised groups of internet users are known to submit such notifications (online vigilante groups have in the past played a significant role in flagging online content, thereby initiating website blocking procedures; Daucé et al., 2019).

The Domain Checker available on the Netoscope website warns users about any malicious activity the checked domain name is/was involved in based on Netoscope partners' assessments. In line with the restricted disclosure and anonymised aggregation discussed above, the results received from the Domain Checker do not show which partner flagged the domain name in question nor when this occurred. The Coordination Center representative explained, making information non-traceable was "the main condition at the start of the project." It means that, although the Coordination Center has access to these details, information about partners' involvement is not disclosed, and this lack of transparency extends to all partners in the project. As Yarnykh explained, Kaspersky sends information "like an email" and is not able to trace how it is subsequently processed.

For some of the international partners, the motivations behind joining the project and the content of their contributions appear to be somewhat different. The partnership with FIFA was established in 2018 in the context of the World Cup that Russia hosted. According to FIFA's advisor on brand protection, Aleksei Shvetsov, "FIFA [would] identify and transfer data to Netoscope about domain names used for phishing in the illegal sale of tickets for the World Cup" ("FIFA i 'Netoskop' budut," 2018). The received data would be analysed by "participants of Netoscope" and resources blocked if illegal activities were indeed identified. SURFnet, a cooperative association of Dutch educational and research institutions aimed at the development and procurement of information and communication technology facilities and knowledge sharing, concluded their agreement with Netoscope in 2017. This followed upon initial contact between SURFnet and Technical Center "Internet" at the Internet Engineering Task Force meeting in Berlin in 2016 (interview with Roland van Rijswijk-Deij). SURFnet had an interest in obtaining access to data on the Russian ccTLDs as part of a larger open intelligence project. Following a year-long negotiation process, an agreement to this effect was signed with Netoscope, on condition that SURFnet reports any relevant threats it finds on the basis of the shared data with its Russian partner. According to Roland van Rijswijk-Deij, SURFnet contributed to the Netoscope database on a single occasion (spam detection) and did not receive information on how their notification was handled.

5.3. Discussion: Implications and Possibilities for Misuse

Yarnykh positively assessed the results of Netoscope since the Coordination Center managed to consolidate collaboration among the leading Russian internet companies in the project. Therefore, he considered Netoscope as "a valuable example to also be emulated on an international level, provided the level of trust, responsibility, and coordination is sufficient to use such a cooperation for the sake of internet stability." Yet, from an internet governance perspective, the project also creates a fundamental vulnerability, especially given the current politicisation of internet infrastructure in Russia, that is of relevance beyond our case. Our study shows that the Coordination Center indeed trusts its partners' assessments and does not check whether information sent to the database is correct. On the other hand, Kaspersky (and presumably, other partners) trust the Coordination

Center and cannot trace how the information they provide is processed by Netoscope. The Netoscope database is non-transparent for all but the Coordination Center and it is precisely this condition of non-transparency that served as the basis for establishing and preserving trust within the project. However, the same condition of non-transparency gives rise to concern related to how the database is/may be used by various end-users and the lack of any (legal) redress for domain name owners. Combined with the lack of verification mechanisms (except for domains flagged by internet users) it risks the trust in it being violated by malicious flagging, i.e., an innocent domain name being accused of containing malware by (an employee of) one of the partners or a targeted website being accused of intentional involvement in a (simulated) attack in order for it to be included in the Netoscope database.

In addition to the fact that most partners are either de facto controlled by the state or have had their independence from the state questioned, a particular area of concern is the lack of information on how Roskomnadzor, as the federal agency involved in executing (restrictive) internet regulation, contributes to the project. While there is currently no evidence suggesting that Roskomnadzor uses the Netoscope database to flag unwanted speech as well as malware (which would negatively affect the reputation of the domain name, which could affect its indexation and recommendation) the governance structure of the project, in as far as we were able to confirm, does not have safeguards against such misuse. Within its current scope of competence, Roskomnadzor may then use Netoscope as an implementation tool, instead of, or alongside the other means of enforcement at its disposal (legal action, fines, preemptive website blocking); although, again, their willingness to do so may only be assumed since, as of yet, no proof of its misuse is available. In such a case, using the governance particularities of Netoscope and the competencies of the partners involved (representing leading search engines, news aggregators, and social media platforms) may prove quite effective in extending internet control mechanisms to the level of DNS infrastructure. Similar to other algorithm-driven forms of hidden censorship (Makhortykh & Bastian, 2020), detecting and exposing such misuse is difficult; the lack of transparency and accountability limits possibilities for exposing misuse while trust in the (abused) system is continually reinforced through its usage. Given that Roskomnadzor did not respond to our interview request, information on its role remains limited.

Applying Schwemer's trusted notifier-model to the information we gathered shows Netoscope possesses all of the model's four features: First, Netoscope is based on voluntary arrangements; second, Netoscope partners act as privileged parties with a direct channel, through the database, to the Coordination Center as the intermediary with the capacity to affect the accessibility of flagged content; third, there is no requirement of preliminary judiciary assessment of whether content flagged by the trusted notifiers is indeed illegal (in this case, there is also no safeguard mechanism within Netocope to verify partners' notifications); fourth, Netoscope's non-transparency and the fact that its functioning is not restricted by a clear legal framework challenges the project from the perspective of the rule of law, legal certainty, accountability, right to due process, and freedom of expression. Netoscope appears to function as a black box not only for the public and scholarly community but also for the partners themselves. Since publicly available information suggests that project partners use the Netoscope database as an input for their algorithmic ranking systems, the inclusion of independent news sources may affect their online visibility.

Netoscope's governance structure emerged from the need to create a condition of trust among competitors in order to share data and collectively work towards reducing malware and phishing within the Russian domains. It emerged from the Coordination Center, which, as a technically-oriented non-profit organisation, is influenced by international practices of multistakeholderism in internet governance. Operating through collaboration with security professionals within its partner organisations, their shared understandings of and trust in the reliability of technical expertise provide the basis for the database and its use. However, the introduction of the Russian Internet Sovereignty Act and the (planned) creation of a national DNS are only the most recent signs of a shift from multistakeholderism to a state-centric tendency in Russian internet governance. This politicisation and securitisation of internet infrastructure in Russia mean that the project's neutrality and "technocratic" nature can no longer be assumed. As DeNardis (2014, p. 18) argued some years ago: "Internet governance structures were originally based on familiarity, trust, and expertise and on 'rough consensus and running code.' Things have changed." The fact that the Russian state has become a stakeholder in the Coordination Center is but one indicator of this trend. The lack of transparency—crucial to its involvement of private partners—creates a lack of accountability. Contrary to the procedural requirements and reporting obligations that pertain to, for example, website blocking, a similar degree of transparency is not provided when it comes to the contents and applications of the Netoscope database, which makes it hard to detect whether Netoscope has been used as a tool for online content control. Moreover, those applications of the database that are particularly relevant for indirect media control (algorithmic downranking of flagged domains) are considered company secrets. Recently, transparency concerns have also been expressed regarding website blocking (Stadnik, 2021). As was mentioned above, the Russian Internet Sovereignty Act enables website blocking through state-controlled deep packet inspection filters which may turn it into a black box. In this respect, both cases signal a worrying trend towards rendering

online content governance in Russia less transparent and thereby less accountable.

6. Conclusion

Russia's push towards establishing a "sovereign" internet has garnered international attention in academic, policy, and rights advocacy circles alike. The possible impact of the policy on freedom of expression, among other rights, has been a key concern in these debates, resonating with the earlier concerns about overblocking such as those included in the *Kharitonov v. Russia* (2020) and *OOO Flavus and others v. Russia* (2020) decisions. Scholarship on media control in Russia, however, has yet to fully embrace the importance of internet governance as an enabling or prohibiting factor. Our aim has been to argue for a broadening of how authoritarian control of online media is studied by looking not just at legislation, media ownership, journalistic culture, or self-censorship, but also by critically examining how key technology and internet infrastructure players are involved in internet governance practices that may affect the online dissemination of news and other information. On the example of Netoscope, we argued that the use of a trusted notifier-model, which is currently gaining in popularity as a way to, for example, address online harm within social media, may be vulnerable to manipulation or abuse without effective legal/procedural safeguards and transparency requirements (although, as of yet, there is no evidence of misuse in this particular case). While further research is needed, our findings suggest there are grounds for questioning the general validity of using trust-based models in non-free media systems as they amplify their inherent weaknesses (e.g., limited accountability). To fully grasp the role and impact of such governance practices that exercise control via (physical, algorithmic) internet infrastructures, an analysis of further cases is required. For example, the recent initiative by Yandex to engage selected media and fact-checking organisations as trusted notifiers to counter "fake news" on its personalised content distribution platform (Yandex Zen) illustrates the urgent need to establish an understanding of media control that reflects the complexity of digital ecosystems today. Our analysis of Netoscope underscores the importance of transparency and accountability mechanisms to safeguard against (future) political instrumentalisation of ostensibly technical or specialist collaborations, systems, and governance structures.

Acknowledgments

The authors would like to thank Mykola Makhortykh and two anonymous reviewers for their insightful comments.

Conflict of Interests

The authors declare no conflict of interest.

References

Balkin, J. (2014). Old-school/new-school speech regulation. *Harvard Law Review*, *127*(8), 2296–2342.

Coordination Center. (2012). *About project*. Netoscope. https://netoscope.ru/en/about

Coordination Center. (2013). *Otchet Direktora ANO "Koordinatsionnyi tsentr national'nogo domena seti Internet" A.V. Kolesnikova* [Report by the Director of ANO "Coordination Center of the national domain of the internet" A.V. Kolesnikov]. https://cctld.ru/upload/files/dir_year_report_2012.pdf

Coordination Center. (2015). *Otchet Direktora ANO "Koordinatsionnyi tsentr national'nogo domena seti Internet" A.A. Vorob'eva* [Report by the Director of ANO "Coordination Center of the national domain of the internet" A.A. Vorob'ev]. https://cctld.ru/upload/files/dir_year_report_2014.pdf

Coordination Center. (2016). *Soglashenie o sotrudnichestve v sfere protivodeistviia rasprostraneniiu v seti Internet informatsii, priznannoi zapreshchennoi k rasprostraneniiu na territorii Rossiiskoi Federatsii, i rasprostraniaiushcheisia s narusheniem zakonodatel'stva Rossiiskoi Federatsii* [Agreement on scientific-technical cooperation in the sphere of counteracting the dissemination of information that is illegal to disseminate in the Russian Federation on the Internet]. https://cctld.ru/files/news/rkn_agreement.pdf

Coordination Center. (2017). *Otchet Direktora ANO "Koordinatsionnyi tsentr national'nogo domena seti Internet" A.A. Vorob'eva* [Report by the Director of ANO "Coordination Center of the national domain of the internet" A.A. Vorob'ev]. https://cctld.ru/upload/files/dir_year_report_2016.pdf

Coordination Center. (2020). *About the center*. https://cctld.ru/en/about

Daucé, F., & Loveluck, B. (2021). Codes of conduct for algorithmic news recommendation: The Yandex.News controversy in Russia. *First Monday*, *26*(5). https://doi.org/10.5210/fm.v26i5.11708

Daucé, F., Loveluck, B., Ostromooukhova, B., & Zaytseva, A. (2019). From citizen investigators to cyber patrols: Volunteer internet regulation in Russia. *Laboratorium: Russian Review of Social Research*, *11*(3), 46–70.

Daucé, F., & Musiani, F. (Eds.). (2021). Infrastructure-embedded control, circumvention, and sovereignty in the Russian internet. *First Monday*, *26*(5).

Deibert, R., & Rohozinski, R. (2010). Control and subversion in Russian cyberspace. In R. Deibert, J. Palfrey, R. Rohozinski, & J. Zittrain (Eds.), *Access controlled: The shaping of power, rights, and rule in cyberspace* (pp. 15–34). MIT Press.

DeNardis, L. (2014). *The global war for internet governance*. Yale University Press.

Drake, W., Cerf, V., & Kleinwächter, W. (Eds.). (2016). *Internet fragmentation: An overview*. World Economic Forum.

Dupont, B. (2017). Bots, cops, and corporations: On the limits of enforcement and the promise of polycentric regulation as a way to control large-scale cybercrime. *Crime Law Soc Change*, *67*, 97–116. https://doi.org/10.1007/s10611-016-9649-z

Dutton, W. (2015). *Multistakeholder internet governance?* World Bank. https://pubdocs.worldbank.org/en/591571452529901419/WDR16-BP-Multistakeholder-Dutton.pdf

Ermoshina, K., & Musiani, F. (2021). The Telegram ban: How censorship "made in Russia" faces a global internet. *First Monday*, *26*(5). https://doi.org/10.5210/fm.v26i5.11704

European Commission. (2017, September 28). *Security Union: Commission steps up efforts to tackle illegal content online* [Press Release]. https://ec.europa.eu/commission/presscorner/detail/en/IP_17_3493

FIFA i "Netoskop" budut borot'sia s moshennichestvom pri prodazhe biletov na ChM-2018 [FIFA and "Netoscope" will fight against fraud in ticket sales for the 2018 World Cup]. (2018, February 13). *RIA Novosti*.

Hill, J. (2012). A Balkanized internet? The uncertain future of global internet standards. *Georgetown Journal of International Affairs*, *2012*, 49–58.

Klein, H. (2002). ICANN and internet governance: Leveraging technical coordination to realize global public policy. *The Information Society*, *18*(3), 193–207.

Kravets, D., & Toepfl, F. (2021). Gauging reference and source bias over time: How Russia's partially state-controlled search engine Yandex mediated an anti-regime protest event. *Information, Communication & Society*. Advance online publication. https://doi.org/10.1080/1369118X.2021.1933563

Kudriavtseva, V. (2020, February 26). Kak ne popast' v seti internet-moshennikov? [How not to get caught in the net of internet scammers]. *Telekanal Kul'tura*.

Litvinenko, A., & Toepfl, F. (2019). The "gardening" of an authoritarian public at large: How Russia's ruling elites transformed the country's media landscape after the 2011–2012 protests for fair elections. *Publizistik*, *64*(2), 225–240.

Lokot, T. (2018). Be safe or be seen? How Russian activists negotiate visibility and security in online resistance practices. *Surveillance & Society*, *16*(3), 332–346.

Lonkila, M., Shpakovskaya, L., & Torchinsky, P. (2020). The occupation of Runet? The tightening state regulation of the Russian-language section of the Internet. In M. Wijermars & K. Lehtisaari (Eds.), *Freedom of expression in Russia's new mediasphere* (pp. 17–38). Routledge.

Makhortykh, M., & Bastian, M. (2020). Personalizing the war: Perspectives for the adoption of news recommendation algorithms in the media coverage of the conflict in Eastern Ukraine. *Media, War & Conflict*. Advance online publication. https://doi.org/10.1177/1750635220906254

Möllers, N. (2021). Making digital territory: Cybersecurity, techno-nationalism, and the moral boundaries of the state. *Science, Technology, & Human Values*, *46*(1), 112–138.

Mueller, M. (2010). *Network and states: The global politics of internet governance*. MIT Press.

Musiani, F., Cogburn, D. L., DeNardis, L., & Levinson, N. S. (Eds.). (2016). *The turn to infrastructure in internet governance*. Palgrave Macmillan.

Napoli, P. (2015). Social media and the public interest: Governance of news platforms in the realm of individual and algorithmic gatekeepers. *Telecommunications Policy*, *39*(9), 751–760.

Netoscope. (2020). *Otchet proekta Netoskop za 4 kvartal 2020 goda* [Report of the Netoscope project on Q4 2020]. https://netoscope.ru/upload/stats/Netoscop_2020_4.pdf

Netoscope. (2021). *Domain checker*. https://netoscope.ru/en/check/?q=

OOO Flavus and others v. Russia, Nos. 12468/15, 23489/15, and 19074/16 (2020).

Poiskovik "Sputnik" prekratil rabotu [Search engine "Sputnik" ceased its operations]. (2020, September 8). *RIA Novosti*. https://ria.ru/20200908/sputnik-1576905844.html

Schwemer, S. (2018). On domain registries and unlawful website content: Shifts in intermediaries' role in light of unlawful content or just another brick in the wall? *International Journal of Law and Information Technology*, *26*(4), 273–293. https://doi.org/10.1093/ijlit/eay012

Schwemer, S. (2019). Trusted notifiers and the privatization of online enforcement. *Computer Law & Security Review*, *35*(6). https://doi.org/10.1016/j.clsr.2019.105339

Sherstoboeva, E. (2020). Regulation of online freedom of expression in Russia in the context of the Council of Europe standards. In S. Davydov (Ed.), *Internet in Russia: A study of the RuNet and its impact on social life* (pp. 83–100). Springer.

Sivetc, L. (2020). The blacklisting mechanism: New-school regulation of online expression and its technological challenges. In M. Wijermars & K. Lehtisaari (Eds.), *Freedom of expression in Russia's new mediashere* (pp. 39–56). Routledge.

Sivetc, L. (2021). Controlling free expression "by infrastructure" in the Russian internet: The consequences of RuNet sovereignization. *First Monday*, *26*(5). https://doi.org/10.5210/fm.v26i5.11698

"Sputnik" iskliuchaet iz poiskovoi vydachi saity, popavshie v "chernye spiski" Roskomnadzora ["Sputnik" excludes sites that are included in Roskomnadzor's "black lists" from its search results]. (2014, September 25). *SearchEngines.ru*.

Stadnik, I. (2021). Control by infrastructure: Political ambitions meet technical implementations in RuNet. *First Monday*, *26*(5). https://doi.org/10.5210/fm.v26i5.11693

van Dijck, J. (2020). Seeing the forest for the trees:

Visualizing platformization and its governance. *New Media & Society*, *23*(9), 2801–2819. https://doi.org/10.1177/1461444820940293

Vendil Pallin, C. (2017). Internet control through ownership: The case of Russia. *Post-Soviet Affairs*, *33*(1), 16–33.

Vladimir Kharitonov v. Russia, No. 10795/14 (2020).

Wijermars, M. (2021). Russia's law "On news aggregators": Control the news feed, control the news? *Journalism*. Advance online publication. https://doi.org/10.1177/1464884921990917

Wijermars, M., & Lehtisaari, K. (2020). Introduction: Freedom of expression in Russia's new mediasphere. In M. Wijermars & K. Lehtisaari (Eds.), *Freedom of expression in Russia's new mediashere* (pp. 1–14). Routledge.

About the Authors

Liudmila Sivetc is a lawyer and a doctoral candidate at the Faculty of Law, University of Turku. Her research interest is connected to internet regulation and freedom of expression. Her most recent work has appeared in *First Monday* as part of the Special Issue *Infrastructure-Embedded Control, Circumvention and Sovereignty in the Russian Internet* edited by Daucé and Musiani (2021).

Mariëlle Wijermars (PhD) is an assistant professor in cyber-security and politics at Maastricht University. She conducts research on algorithmic governance, media freedom, and the human rights implications of internet policy. She is co-editor of *The Palgrave Handbook of Digital Russia Studies* and *Freedom of Expression in Russia's New Mediasphere*.

COGITATIO

Media and Communication (ISSN: 2183–2439)
2021, Volume 9, Issue 4, Pages 39–51
https://doi.org/10.17645/mac.v9i4.4213

Article

Trolls, Pressure, and Agenda: The Discursive Fight on Twitter in Turkey

Uğur Baloğlu

Faculty of Applied Science, Istanbul Gelisim University, Turkey; E-Mail: ubaloglu@gelisim.edu.tr

Submitted: 24 February 2021 | Accepted: 23 July 2021 | Published: 21 October 2021

Abstract
Censorship, banning, and imprisonment are different methods used to suppress dissenting voices in traditional media and have now evolved into a new form with bot and troll accounts in the digital media age in Turkey. Is it possible to construct a bloc with counter-trolls against the escalating political pressure on the media in the post-truth era? Are counter-trolls capable of setting the agenda? This article discusses the possibility of constructing a bloc against the escalating political pressure in Turkey on the media through counter-trolls in the context of communicative rationality. First, it observes the ruling party's troll politics strategy on Twitter, then examines the counter-discourses against political pressure today; thereafter it analyzes the discourse in hashtags on the agenda of the Boğaziçi University protests. Firstly, 18,000 tweets are examined to understand the suppress-communication strategy of the AK Party trolls. Secondly, the agenda-setting capacity of counter-trolls is observed between January 1, 2020, and February 5, 2021, and 18,000 tweets regarding Boğaziçi protests are examined to analyze the communication strategy of the counter-trolls. The study shows that the populist government instrumentalizes communication in social media, and Twitter does not have enough potential for the Gramscian counter-hegemony, but the organized actions and discourses have the potential to create public opinion.

Keywords
agenda; civil society; communication strategy; counter-trolls; populism; trolls; troll politics

Issue
This article is part of the issue "Media Control Revisited: Challenges, Bottom-Up Resistance and Agency in the Digital Age" edited by Olga Dovbysh (University of Helsinki, Finland) and Esther Somfalvy (Research Centre for East European Studies at the University of Bremen, Germany).

1. Introduction

Today, 4,5 billion people use the internet on earth, where approximately 7,8 billion people live, and 3,8 billion of these users have social media accounts (Kemp, 2020). This shows us that more than half of the world population is online. When we narrow the perspective and focus on Turkey, we can observe that 54 million of the approximately 83 million people (Turkish Statistical Institute, 2020) have social media accounts and this rate is parallel to developments in the world, namely, about 60% of the population uses social media (Kemp, 2020). The increasing number of users of social media indicates that our communication practices will be different from the past. As a matter of fact, we have been experiencing a similar change in communication practices of the typographical age that McLuhan mentioned in *Gutenberg Galaxy* (1963). This also makes us question the relation between one's connection to the reason. Then, in Western notion, men who have the ability to illuminate the darkness as "homo rationalis" (Çiğdem, 2004, p. 55) turn into "homo irrationalis" in today's post-truth discussions (Fasce, 2020; Levitin, 2017, p. 2; Pinker, 2018, p. 371).

Emphasizing the normativeness of the concept in the post-truth age in which new communication practices are experienced, McIntyre (2018, p. 6) indicates that truth is subordinated to the political one and thus contextualized within the framework of its own ideological perspective. This period, during which the truth is deformed, describes an environment in which objective (rational) phenomena are abandoned and non-objective (irrational) personal opinions are dominant. As a matter of fact, Keyes (2004) also points out that during this

period, lying penetrated into new communication practices, making it easier to deceive people. This can cause different problems in the context of interpersonal communication, but it can lead to greater problems in the channels where news circulation is provided on issues of public concern. The information circulated through social media tools is not subject to any editorial process and has not been inspected, which provides a separate dimension to fake news discussions. In addition to its alternative and liberating potential, these platforms, where polarizing and otherizing discourses are easily got into circulation, pave the way for an environment which is suitable for authoritarian-populist politicians (Grinberg et al., 2019). Recent research on fake news on social media platforms (especially Twitter, due to its alternative news media feature; Kwak et al., 2010) that allows a person to hide behind an anonymous identity also supports this (Bovet & Makse, 2019; Brummette et al., 2018; Recuero & Gruzd, 2019). So, what kind of actions does the political authority take if it wants to consolidate its power in a media environment where the truth is ambiguous? How does political power respond to the liberating potential of new communication platforms in an environment where traditional media is neutralized in the context of media, politics, and the intricate relationship of capital? Such questions are important for understanding the way right-wing populist political authorities communicate through the media. Media has been one of the most transformed "power" in the single-party government process since 2002 (Çam & Yüksel, 2015). The transition (flow) of ownership structure to conservative capital which is close to political authority is one of the main indicators of ideological monologism in traditional media (Media Ownership Monitor Turkey, 2019). The colonization of the media by the party, conceptualized by Bajomi-Lázár (2013) through Hungary, points to a similar process in Turkey. During the 2013 Gezi Park protests, the economic and ideological hegemony of the media controlled by political authority was surpassed, and social media was then used as an alternative means of communication. This determines the attitude that the political authority will introduce to social media in the future.

The relationship between power, capital, and the media in Turkey indicates a long-term process. Today, however, as a result of restrictions, bans, and obstructions on freedom of expression, the pressure on the media has augmented. This is confirmed by the World Press Freedom Index of Reporters Without Borders (2020), with Turkey ranked 154th out of 180 countries. Today, political power does not merely use the means of repression, such as prohibition, obstruction, or investigation and imprisonment of journalists (The Journalists' Union of Turkey, 2021); it develops different tools to make media domination invisible and expand the field of discourse. Troll-politics, which describes changing the agenda in the political context, making propaganda, producing disinformation, and/or producing hate speech to suppress opposing/oppositional view, is perhaps the most serious technique in these tools. The fact that social media is a suitable platform for producing user-derived content helps the political authority, through trolls, to manipulate people in order to lure them into futile disputes or deviate them from provocatively getting involved in an ongoing discussion. That is to establish both the dominance of its own discourse and to divert opponents from the subject by decontextualizing the discussions (Binark et al., 2015, pp. 127–128). The way such political discussions are held in channels such as Twitter is significant in terms of projecting the image of supporting freedom of expression. This new communicative strategy creates the illusion that propagandist statements, lies, and polarizing discourses circulate through the people, not directly from the political power, causing the truth to be ambiguous. The effort of authoritarian-populist political authority to create/set its own agenda through trolls by making insignificant the truth or distorting it out of discourse in the post-truth period opens up a space for it to reproduce its hegemony.

Since 2013, AK Party has been actively displaying a policy on social media through its digital office headquarters (Altuntaş, 2015). AK Party's troll army's attempt to put pressure on Twitter is important both to figure out its populist politics and to determine a strategy against it. Since the relationship between the threat of liberal democracies turning towards authoritarianism and the negative trend in the communication paradigm necessitates organized struggle against authority on social media platforms. The article seeks to answer the question of how to create a bloc against the troll policy of political authority in the post-truth age. Besides, it asks if the counter-troll politics have brought the common good rather than polarizing society. The study will first examine the counter discoursive attacks of groups trying to make their voice heard against political authority today after observing AK Party's troll politics strategy via Twitter (Bulut & Yörük, 2017; Karatas & Saka, 2017).

2. Data and Method

In this research, two different methods were used, namely, content analysis of hashtags and thematic analysis of trolls' tweets. Content analysis is a quantitative approach that summarizes the numerical outputs of variables, whereas thematic analysis is a qualitative method that emphasizes "constellations" by examining the patterns of meaning in texts (Neuendorf, 2019, p. 213). It was observed how long the hashtags remained on the agenda and the number of retweets to reach the most influential content in the data collective section of the study. TAGS v6.1.9.1 and Twitter Search were used to scan all sets of tweets sent by active users. Categories of Stanford Internet Observatory Cyber Policy Center were considered for the detection of troll accounts. The tweets were examined considering the structural features such as the joining date, account name, profile

information and photo, consistent political tweets, and being an active user (tweeting per day; Fuchs, 2017, p. 54; Grossman et al., 2020). The tweets were selected from the users with the most followers, most retweeted, and likes among 18,000 tweets. Besides, the content of the tweet threads was also checked in accordance with the theme. The network analysis of Hafıza Kolektifi (" 'Ak Trol'lerin haritası," 2015) on Twitter demonstrated that Aktrolls, namely, trolls supporting AK Party and AK Party executives were in contact. In the determination of troll accounts, the accounts associated with the political party-related, politicians, and the relationship they establish with their own troll groups were also taken into account in terms of embodying the relationship with the political party (retweeting and/or liking each other's tweets). Based on the information presented in the first analysis, tweets posted in hashtags that were trending at the Boğaziçi University protests—which started with students and academics protesting the appointment of a new rector in a presidential decree—made up the sample group.

First, their posts were divided into themes to analyze the communication strategies of the trolls supporting AK Party. In the case study, the discourses of 18,000 random tweets posted with the hashtag #DevletiminYanındayım (IamWithMyGovernment) and the tweets of 19 active accounts with the most followers among 18,000 tweets were examined. Since the communication strategies of accounts with troll-politics contain similar themes, the analysis is limited to 19 accounts. #Kabekutsalımızdır (KaabaIsOurHoly), #KabeyiSavunanFişleniyor (TheOnesDefendingKaabaAre Blacklisted), and #ProvokatörlerdenİşgalGirişimi (InvasionAttemptByProvokers) hashtags were included in the study because they were thought of as launching flares for the Boğaziçi protests.

In the second analysis, hashtags used by the ones aiming to produce counter-discourse and their capacity to create an agenda, and how many people they reached, were examined. Within the framework of the discourses developed against political power, the selection of hashtags was limited to topics that could reach a minimum of 100,000 tweets and be on the agenda between January 1, 2020, and February 5, 2021. The maximum number of tweets and how long the hashtags remain on the agenda data were found by examining daily Twitter trending data. Then, the "troll-politics" strategy of opposition groups aiming to establish counter-hegemony by producing counter-discourse was examined. The author tried to find out whether the counter-discourses were produced by trolls. Their posts were divided into themes to analyze counter-trolls' communication strategy. In the case study, the discourses of 18,000 random tweets posted with the hashtag #AşağıBakmayacağız (WeWillNotLookDown) were examined.

Finally, the discourses of the second group and the ability to create an agenda on Twitter were interpreted in the context of communicative rationality and counter-hegemony in response to the data in the first group. The capacity of troll-politics to form a counter-bloc in the context of communicative rationality was discussed within the framework of the analyzed data.

3. Media, Populism, and Troll Politics

On July 1, 2020, President Erdogan said, "we want these social media platforms completely shut or controlled after bringing the issue to our parliament" (Erem, 2020, para. 5). Every political authority needs the support and control of the media to reproduce its own rulership. AK Party's policies on the media over the past 18 years show that it has continuously intervened in the media to maintain its own discursive order (Erem, 2020). In order for the political authority to convey its populist policies to the public within the framework of its own reality, it is essential that the media be under its own control. For this reason, the decrease in public consent, especially in times of economic crisis, features the oppressive aspect of the government, as seen in Erdogan's statement (Gramsci, 1971, p. 246). The media policy, which began with the seizure of Star TV (which belongs to Uzan Group) in 2004 as soon as it came to power (Duran, 2015, pp. 20–21), continues with the transfer of the media sector organizations of Doğan Group, which is called the "flagship" of the press, to Demirören Group in 2018. AK Party tries to actualize not only an economic but also an ideological, social, and cultural transformation while creating its own hegemony in the field of media. While the discussions on how successful the government is in this regard continue, a striking point is the polarization of society (Bulut & Yörük, 2017). The fact that populist politicians try to establish their own discursive dominance by polarizing society is shaped by two different lifestyles created between the public and the elite. The populist identity that stands out as authorized to protect the interests of the people promises to transform the elitist institutions that existed in the past (Aytaç et al., 2021, pp. 5–6). Turkey's modernization process is different from the West. It is crucial to figure out the success of populism today to understand the reactions of the masses that are not represented in the public field in Turkey. The policy is to plan a new path by keeping secularism and Islamic thought in a certain balance policy (Kaya, 2015). Likewise, the ruralization of the urbanites with the migration from the village to the city has progressed faster than the urbanization of the villagers. This has de-eliteized Turkish modernization. Especially after 1950 (after the Democratic party came to power), Turkey began becoming Anatolianized (provincialized; Mardin, 1991, p. 276). The populist politician who shows himself as an anti-elitist and anti-intellectual (Ghergina et al., 2013, p. 3) aims to create the image of a global political leader with their nostalgic bond with the past and heroism anecdotes. At this point, he reproduces his discourses in a populist context, based on the power of the frantic masses, which has been repeatedly excluded from

the public sphere in the past. The media, on the other hand, uses different methods such as "claim" and "repetition" (Le Bon, 1997, p. 47) to internalize some thoughts and beliefs of the masses and consolidate the power of political authority.

2013 points out a break in Turkish media history. People tend to social media as a result of the unsuccessful portrait of traditional media during the Gezi Park protests. Thus, political reporting reflex develops on social media, especially on Twitter (Karatas & Saka, 2017, p. 385). In the light of these developments, despite the decentralized nature of social media, the fact that it is an organized communication platform against the power makes AK Party give more importance to online activity. Authoritarian-populist governments, which use similar troll armies as propaganda vehicles in the world, manipulate the public in the context of their ideology with the information they circulate through platforms such as Twitter. Today, the spread of disinformation takes place not through old mass media such as radio and television, but rather with tweets, bots, and fake social media groups (Weedon et al., 2017).

Each period creates its own discourse and concepts. The concept of trolls also indicates a popular concept thanks to new communication technologies. We can get a clearer grasp of troll-politics when we consider that the AK Party has instrumentalized communication practices in order to establish its hegemony and expand its power.

3.1. Analysis

Twitter, users of which are increasing day by day and gaining alternative news media features, has become a platform where AK Party institutionalizes political communication negatively in terms of setting, changing or decontextualizing the agenda, polarizing society, spreading propaganda, and producing hate speech (Binark et al., 2015, p. 128). The accounts and tweets reviewed in the analysis indicate that the discourses that develop against the Boğaziçi protests are categorized within three different themes. The first is the polarizing discourses in which the opposition of "me and the other" is highlighted by referring to nationalist and/or religious elements and identity politics is at the center. Laclau (2005) notes that collective identity and the other positioned in front of it are the main means of understanding populism today. We see a similar political discourse in the tweets of trolls on the "us" and "them" polarization of political power.

As part of the Boğaziçi protests, some students held an exhibition on campus. One of the images in the exhibition was a painting standing on the ground with the Kaaba figure on it. There was a Shahmaran figure (a mythological creature in Anatolia that is believed to bring abundance, wealth, happiness, luck, and protect people from evil eye, has been the subject of legends, with a head in the form of a human and a body in the form of a snake) in the middle of the painting, and LGBTI+, lesbian, trans, and asexual flags were placed at its four corners:

> @TheLaikYobaz: Against the cowardly children of Mount Olympus, we are the brave sons of Mount Hira! (SON LAİK BÜKÜCÜ, 2021a, translation by the author)

> @KacSaatOlduTR: Why are you silent, Mr. Kemal? Aren't you going to say anything against provocateurs who have threaten promising college students, targeted the Kaaba, represented the people of Lot? Or are you an enemy of religion, too? (Kaç Saat Oldu?, 2021b, translation by the author)

The picture of the Kaaba figure surrounded by LGBTI+ flags at Boğaziçi University was criticized by AK Party supporters for insulting religious values. The picture, described as a derogatory attitude towards sacred values, provided the AK Party with the illegitimate populist way of communicating towards polarization within the framework of religious discourse. As it can be seen in the tweets above, a Muslim/non-Muslim contrast is created by referring to religious elements. In the first tweet, the enmity of Greek mythology and Islam and Greek-Turk are based on religious-nationalist discourse, while the second tweet creates a polarization over religious hostility by targeting the opposition party. In the tweets reviewed, it was observed that the populist communication strategy, in which polarizing discourses are seen commonly within the framework of similar identity politics, was implemented. Within the framework of me–other polarization, "me" points out the people with a statist and Muslim identity, while the "other" represents the people opposed to Islamic values and the government.

The second one is manipulative tweets containing threats, profanity, humiliation, and/or disinformation involving political, social, or direct personalized discourses. Özsoy (2015), who conducted a study on trolls in Turkey, states that trolls generally prefer a provocative, manipulative, and negative language of discourse (p. 537). When the troll-politics accounts were examined, it was found that a similar language of communication prevailed. Aktroll's follows a strategy that constantly emphasizes the distinction between "me and the other" and constantly targets the other, and in this context, re-creates the issue or the agenda within the framework of its own reality. Thus, the subject or agenda is distorted from its own context and tried to be cut off in the abundance of different discussions:

> @THEMARGİNALE: There you are, we said before. Their problems are neither rector nor university. They're just trying to ignite the fuse of a new attempt. (Marginale, 2021, translation by the author)

> @emirbereket: WE DO NOT WANT TERRORISM on the street, in universities, in public, in parliament. (Eemir Bereket, 2021, translation by the author)

In all the videos and images in the hashtags examined in the Boğaziçi protests, protesters are shown in an

offensive and damaging context. At this point, reality becomes a phenomenon that varies from individuals' point of view. However, this includes the threat of suppressing the voices of opposing views and the danger of preventing democratic participation. That's why, "unlike trolling, troll-politics is serious," as Merrin (2019, p. 291) emphasizes. As it is seen in the above tweets, trolls are trying to manipulate AK Party supporters into believing that the protesters are terrorists by invoking the July 15 coup attempt. Most manipulative tweets consist of offensive discourses directly targeting the opposition wing with negative communication styles such as hate speech and humiliation. Likewise, research on troll culture in Turkey demonstrates that Aktrolls follow a similar communication strategy (Binark et al., 2015, p. 138). In addition to the research findings, it was observed that within a year, the manipulation was created through the posts circulated in a language that was condescending, incriminating, and derogatory especially for Ataturk and Republican People's Party (CHP), and that otherizing language such as #AtatürkDiktatördür (AtaturkIsDictator), #FetöcüKemal (FetöMemberKemal), #KemalizmYıkılıyor (KemalizmAreBreakingDown), and RezilsinCHP (DisgracefulCHP) was used intensively.

Thirdly, there are the conspiracy theories frequently used by right-wing populism, often pointing to the ambiguity and popular theories to minimize the consequences of unexpected effects in moments of crisis (Wodak, 2015). Although disinformation and conspiracy theories did not emerge with social media, its circulation became widespread and its quantity augmented. The common point in the analyzed tweets is the emphasis that the protests are led by different enemies. The tweets underline that protests are planned by people or groups that are different from each other but intersect at some points such as Soros, USA, Mason organizations, and Fetö:

> @AntepliMamato: The Bogazici was saved from the "Colonial" occupation. -The military was saved from the "Feto" invasion. -The mountains were saved from the "PKK" occupation. -Foreign Affairs was saved from the "*Monchère*" invasion. -The bureaucracy was saved from the "Mason" occupation. -The industry was saved from the 'Bourgeoisie' occupation. That is because they are going mad. (Mamo Dayı, 2021)

When past network maps of trolls are examined, it shows that they are working organized to silent opponent voices and create a lynching environment in a social media environment (Bulut & Yörük, 2017; Karatas & Saka, 2017). It is possible to say that AK Party's own troll army is working more organized and coordinated than they have done in the past. Furthermore, it can be said that it is not only to suppress opponent voices and manipulate them, but also to take the existing agenda out of context and use troll-politics as instruments by changing the direction of the debate. In a year studied, Aktrolls constantly try to respond to counter-trolls with counter-attacks. For instance, it always tries to change the trends with populist and accusatory responses like #geziihanettir (GeziIsBetrayal) and #GeziDarbesi (GeziCoup) after a few hours after hashtag #GeziyiSavunuyoruz (WeDefend Gezi) was trending, or #Bismillah, #ErdoğanınYanındayım (IamWithErdogan), and #ŞehitlerTepesiBoşDeğil (MartyrsHillIsNotBare) after #negülüyorsunerdoğan (WhyAreYouLaughingErdogan) hashtag was trending. It is also a remarkable detail in which different troll identities take part in different tasks. While some trolls work only to increase the number of retweets to support the hashtag, some trolls produce conspiracy theories. Trolls, who are also in touch with real users (e.g., @elonue, @zekibahce, and @ERKANTAN__), try to mobilize masses by consolidating the undecided audience.

4. Democracy, Civil Society and Counter-Troll Politics

Although the intricate relationship of traditional media with capital and political authority imposes an ideological and technological simplex communication, ideas have been put forward that those social media platforms have evolved towards a participatory, interactive, and collective system and that all these developments strengthen democracy and save it from centralized domination (Hermida, 2010; Jha & Kodila-Tedika, 2020). Is this practically possible for countries like Turkey that have still not created a culture of democracy (Karpat, 2010, p. 53)? Turkey's transition to democracy was not actualized by the institutionalization of civil society, but by the efforts of the state elites. In other words, democracy in Turkey has not developed due to the lack of a civil society notion—inherited from the Ottoman Empire, where political, economic, and social power was gathered at a single point (Heper, 2000, p. 78). Likewise, when Turkey's struggle for democracy is examined in a historical context, it points out to a process that has been constantly interrupted by forces such as the army and political authority rather than the demands of the people. One of the most important reasons for the inability to institutionalize democracy is that Turkey's industrialization process starts late and a large part of society lives in villages (Kongar, 2001). When this is evaluated within the framework of urbanized/non-urbanized societies, cultural differences between NGO's and religious-ethnic communities can be observed more clearly.

The deep relationship of the culture of liberal democracy with civil society can help us understand today's communication practices (Fukuyama, 1995). In this context, the idea of Habermasian civil social power has the potential to gain its own legitimacy against today's authoritarian power as a communicative power. However, it can be said that this is possible in an organized and systemic way with the idea of a civil society institutionalized by getting rid of the domination of state and political parties as an initiator of the opinion leaders of society as expressed by Göle (2000). Because civil

society should not be positioned as an over-politicized anti-government directly in the face of political authority. Thus, it contributes to the culture of democracy with civil society, which stands out from the oppression of authoritarian politics and creates an autonomous field for itself. This can only be achieved through "the transition from identity politics to interaction policies" (Göle, 2000, pp. 80–81).

Despite the post-republican Westernization moves in Turkey, the state is positioned as an absolute device of hegemony, where an East-type tradition still brings out itself in political, social, and cultural reflexes because of its connection with the past. In this context, depoliticizing the public sphere does not allow the implementation of a strong interaction policy. Overcoming this situation—in other words, the non-feudalization of the public sphere—can be achieved by the politicization of social life, the rise of citizen journalism, and the struggle for freedom of expression (Habermas, 2013, p. 17). The ideal Habermasian communication environment conceives an environment in which dissident people interact with each other and carry out their own ideas freely, without exclusion and within the framework of the politics of respect. According to Habermas, people/groups in social practices with communicative rationality can achieve common good through collective actions. At this point, communicative action is the dominant element of participatory democracy. By creating an environment where individuals can freely defend their own ideas/arguments in ideal discussion environments, interaction is provided with the exchange of opposing views (Habermas, 2001). Whether Twitter and/or other social media platforms provide such an environment by democratizing communication is still a matter of debate. So, can Twitter allow counter-troll groups to build counter-hegemony in a Gramscian sense? According to Gramsci, hegemony is:

> Made possible by the dialectic togetherness of force and consent; accordingly, the ruling class must have ideological and institutional foundations along with material forces in order to achieve the consent it needs beyond a difficult domination and to build its hegemony. (Akgemici, 2019, para. 2)

4.1. Analysis

In a study on Twitter, it is found that 73% of trend topics come up only once, and 31% stay on the trends for just one day (Kwak et al., 2010, p. 597). When the duration of counter-discourses and the number of retweets are examined as seen in Figure 1, meaningful parallelism is established with the research of Kwak et al.

The remarkable fact of the study's findings is that the country's agenda and the topics that remain on the trends on Twitter for more than a day are partially the same. In total, only four of the 138 issues remain on the trends for more than a day, and only two of them are on the country's agenda, which indicates the political authority's inaction on the demands of the opposition wing. #CezaevlerindenHaberVar (NewsFromTheJails) and #EbruÖlüyorAcilTahliye (EmergencyEvacuationEbruIsDying), headlines led by the Peoples' Democratic Party, are on Twitter trends, but are not on the country's agenda. However, the topics (#CocukİstismarınınAffıOlamaz [ChildAbuseIsInexcusable] and #AşağıBakmayacağız [WeWillNotLookDown]) that bring together artists, politicians, ordinary people, and anonymous accounts remain on the agenda. When analyzed, collective movements of artists and different political identities are factors in determining the country's agenda.

When examined in the context of the topic as seen in Figure 2, the contents are clustered within five different themes. In 3% of the content, hashtags such as #10Kasım (November10) and #AtatürküÇokSeviyorum (ILoveAtaturkSoMuch) against populist policies conducted through the people and the elite are carried out to remember republican values. Likewise, femicide is a problem that individuals struggle with, regardless of political ideology. Over a year of review, the most recurring contents have been related to officer appointments. Contents related to officer assignments demonstrate that manipulation of bot accounts may be on the trends in a short period of time. The negative correlation between unemployment and employment in Turkey provides an environment for individuals to develop a pragmatic communication language and make their own problems visible.

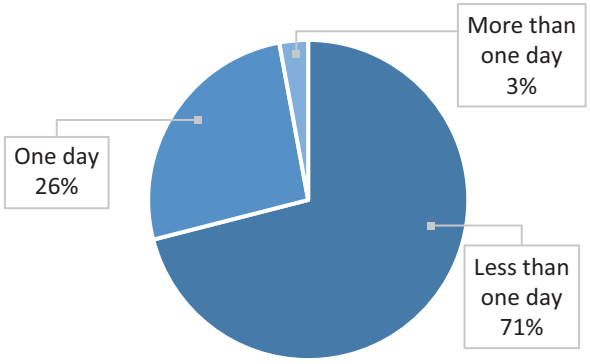

Figure 1. Duration of counter-trends on the agenda.

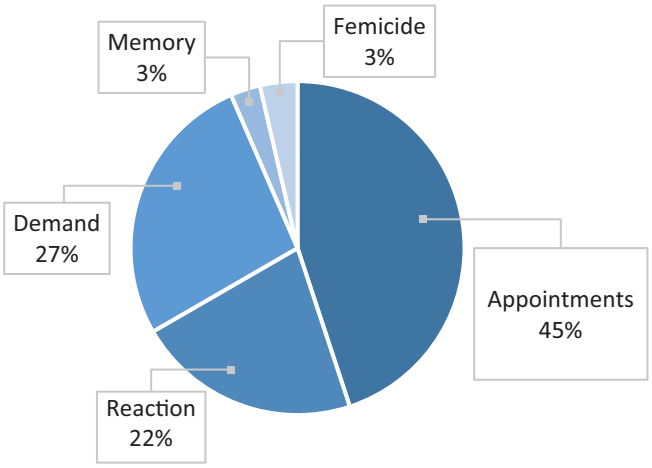

Figure 2. Themes of Twitter counter-trends.

Twenty-seven percent of the analyzed agendas constitute democratic demands from the government. Eleven of the 138 agendas focus on the improvement of sanitary requirements and release of political prisoners in the pandemic; 14 focus on postponing students' exams in the pandemic; and nine focus on improvements in the salary and personal rights of civil servants. These groups use Twitter for pragmatic purposes to attract attention and make their voices heard, even if they are not on the country's agenda. Masses looking for solutions for their own interests instead of achieving common good through communicative rationality turn Twitter into a medium where the populist language is found rather than the Habermasian ideal communication environment. However, this situation is closely related to the non-institutionalization of civil society logic in real social life. As a matter of fact, the individual who cannot be organized in real social life tries to realize his democratic demands indirectly by creating public opinion.

One of the most crucial findings of the study is that there are organized reactions to negative communicative actions of power. "Reaction," which constitutes 22% of the trends, is the counter-discourse formed by meeting with different segments of society on issues such as unemployment, adverse working conditions, pressure on the opposition media, and reaction to social media restrictions. The discourses and interpersonal communication networks examined indicates that there is a disorganized communication on issues where there are no political parties. It can be said that organized communication usually takes place through the efforts of the Republican People's Party and Peoples' Democratic Party. For example, the posts about the termination of the program attended by Istanbul Mayor Ekrem İmamoğlu on CNN Turk one hour early; and the artist Helin Bölek, who died in a hunger strike in protest of Grup Yorum's concert ban. Besides, after Turkish soldiers were killed in Idlib, counter-trolls develop a nationalist discourse in protest of President Erdogan laughing as he got across a memory of Trump in his speech. One of the common features of the discourses developed directly in response to the government is the use of humor.

In Figure 3, President Erdogan is criticized within the framework of his economic action plan in the pandemic—a campaign called Together We Are Enough My Turkey. In the study, it can be said that the people or groups producing counter-discourse conduct partially a troll-politics strategy. In particular, counter-troll-politics stand out on issues such as "government resign," which are likely to be politicized and polarized rather than issues that concern the whole society and are sensitive, such as "censorship law" and "child abuse."

Figure 3. Humorous reaction for measures taken in Turkey against pandemic. Source: conta cunte (2020).

The tweets in the second analysis show that the discourses in the Boğaziçi protests were categorized within five different themes. The structuring of traditional media within the framework of the control of political authority prevents the public from reaching accurate information. In this sense, Twitter becomes a platform where different perspectives and discourses circulate and the public is informed. In the tweets reviewed, detentions, press releases, and police interventions are more likely to be conveyed through various instruments such as writing, photos, and videos. In this sense, most

content is intended to inform the public in the context of the communication network. For instance, press releases and the protests of Boğaziçi University academics got across to the public without interruption (Boğaziçi Dayanışması, 2021). This can be understood as an effort by counter-trolls to expand the information network on Twitter and create public opinion. The second feature of the tweets that develop counter-discourse is that the discourses are associated with Ataturk in order to preserve the values of the republic against neo-Ottoman discourses of the government.

The modernization project in Turkey was carried out with the separation from the Ottoman Empire under the leadership of Ataturk. Today, the AK Party is following a policy that wants to transform the founding values and identity of the republic (Ongur, 2015, p. 416). In this context, counter-trolls reflexively protest against AK Party's counter-policy on Ataturk's founding values, often with content that constantly remembers and quotes Ataturk as seen in Figure 4. Thus, a discourse which came up was not only about Boğaziçi University, but also against social transformation and re-identification policies.

Thirdly, it is observed that the content uses a language similar to the manipulative and polarizing discourses used by the Aktrolls. Tweets in which "me and the other" controversy are produced by counter-trolls and often use derogatory and otherizing language have characteristics similar to the populist communication strategy used by political power:

> @sigaramcamel: The staffed state of ignorance, you are jealous of the children in schools that you would not be able to enroll even if you made a camel picture with your ass. (Kapheros, 2021, translation by the author)

> @MYazar212: Neither can you manage the education! Nor can you manage the health! Nor can you manage the economy! All you can manage are the projects of the partisans. He says, "Whoever has money crosses my bridge." (MYazar212, 2021, translation by the author)

In general, it is observed that an incriminating and furious language against the economic, societal, and educational policies of political authority is present in such polarizing manipulative tweets. In the examples above that, a language that is condescending to political authority is used.

Another communication strategy used by counter-trolls is the humorous discourses on which many studies have been made in the Gezi Park protests. Political humor and intellectual accumulation are considered as a counter-hegemonic strategy developed against the hegemony that the political power is trying to institutionalize (Değer, 2015, p. 319). For example, "@kafayikirdim: Roses are red, Melih is Bulu, He stole an article, And a rectorship, too" (**idilos bébé** ***aka. youroti***, 2021).

As seen in the example, it is analogically described that Boğaziçi University rector Melih Bulu usurped the office of the rectorship suddenly and unexpectedly with plagiarism accusations in his doctoral thesis. A critical language, similar to Rabelais' *Gargantua*, was used: "criticism of the corrupted legal system, the traders who defraud the public and those who exploit religious values, with a humoristic but strong language" (Baloğlu, 2019, p. 224). In this context, humorous language is the basis for breaking the wave of fear created by authoritarian policy.

Finally, it is observed that counter-trolls develop ideal discourses that are structured within the framework of "common good," which does not marginalize, polarize, and does not insult. Such discourses contribute to the institutionalization of interactive democracy with the dialogical character of the communicative reason, as well as play an active role in building the common good. Such discourses, in which the ideal state of speech that Habermas speaks of in communicative actions are created, aim at inclusiveness, not exclusion.

Figure 5, indicating that Boğaziçi University students are against any ban, shows that the group described

Figure 4. Counter-trolls often refer to Atatürk in quarrels. Note: "We learned where to look from our Father" (translation by the author). Source: Kaç Saat Oldu? (2021a).

Figure 5. Boğaziçi students protested the headscarf ban in universities. Source: Yakut #23 (2021).

as the republican elite also protested against the headscarf ban. In this context, it is underlined that the Boğaziçi protests are not an act developed against pure political authority: "@budirenisi: We repeat that our action is peaceful and repeat our demands!" (Boğaziçi Direnişi, 2021).

These statements of the Boğaziçi *Direnişi* (Boğaziçi Resistance) indicate that a language is used that pays regard to the public interest, aims at the common good, and does not marginalize. The Boğaziçi Resistance, organized as a non-profit non-governmental organization, tries to convey the extent of the protests directly without falsifying reality by providing continuous information circulation on Twitter.

5. Discussion and Conclusion

Based on the research findings, it is seen that AK Party instrumentalizes trolling. This is a maneuver in the political communication of the AK Party. Adopting the concept of majoritarian democracy in real social life, the AK Party aims to create social, religious, and moral pressure on social media by trying to institutionalize the tyranny of the majority with its troll army on Twitter. With the bombardment of information by trolls, the circulation of information becomes excessive and reality becomes ambiguous. In natural disasters like earthquakes, Aktrolls change the trends with hashtags such as #DevletMilletininYanında (TheGovernmentIs WithTheNation). The day after an earthquake, the hashtag #DevletimizVarOlsun (LongLiveTheState) is brought to the forefront in İzmir. A similar situation is observed in the Boğaziçi protests. The passive resistance, which became the trend against the appointment of rectors with the hashtag #KabulEtmiyoruzVazgeçmiyoruz (WeDon'tAcceptWeDon'tGiveUp) is drawn into a different context with #KabeKutsalımızdır (KaabaIsOurHoly) counter-attack against the people. The manipulative discourse, that begins with the otherizing and targeting of LGBTI+ individuals, and hate speech are drawn into the context of immorality and devaluation of the sacred value, reflecting the Boğaziçi protests as an act of violence. The most significant point here is that counteractions can be led by manipulations of power. Besides, thoughts/ideas accepted outside the field of hegemonic discourse as if LGBTI+ are prohibited within the framework of freedom of expression. This means that troll-politics disrupts communicative rationality by targeting people or groups. For instance, although it is claimed that there is freedom of expression, hundreds of people are investigated and imprisoned because of their Twitter posts ("Freedom of expression," 2021).

Troll-politics are more about communicative irrationality than communicative rationality. The reconstruction of the lifeworld depends on strengthening the communicative action. For this purpose, the intersubject mind should be prioritized over the subject-centered (Habermas, 2001). However, the discourses developed by Aktrolls via Twitter are aimed at disrupting the communicative action. Almost all of the tweets have features/characteristics that generate discourses supporting the government, polarize, and disrupt public integrity. Thus, communicative irrationality on Twitter disrupts the type of negotiation-oriented action and supports authoritarian governments to take a hegemonic form. It also marginalizes those who defend fundamental human values—such as human rights, justice, and freedom—and directly accuses the protests of doing terrorist activities. This pushes authoritarianism into a concept legitimized by the people, not by the state apparatus. There is a deep relationship between consenting to authoritarianism and selective exposure to information. On platforms like Twitter, people get the information they want or support in echo chambers (Colleoni et al., 2014). The expression of the envisagement of community formed by people with similar thoughts in the news media results in the reinforcement of the values, beliefs, or opinions held. This becomes a communication style that deepens the opposition of "me and the other," results in more polarization of people and provides the fundamental component of today's populist politics. This indicates a similar process for Aktrolls and counter-trolls because in both groups there is information that will justify their beliefs. For instance, videos circulating in hashtags against the Boğaziçi protests show protesters attacking police (SON LAİK BÜKÜCÜ, 2021b), while videos circulating in the hashtags #aşağıbakmayacağız (WeWillNotLookDown) and #boğaziçidireniyor (BogaziciIsStandingOut) show police attacking protesters (Zenibya, 2021).

For political authority, the circulatory channels of information are crucial. Media capture (Schiffrin, 2018) not only explains much about the mass communication of authoritarian regimes, but also gives an idea of the rise of right-wing populism and how political power maintains public control. Then, when information circulation is controlled, communication with the public is carried out in the context of ideological monologism. At this point, it is possible to say that hegemony is changing its form with new communication technologies. It is observed that the manufacturing of consent is achieved in new communication technologies without the need for large media companies, but again by acting rationally with trolls. In this context, digital media is more likely to make media capture and can put legal practices in a legitimate order. These findings are in line with what Schiffrin (2018) has summarized as "rather than disrupting media capture, the digital age in some ways appeared only to change how it is manifested" (p. 1036).

The research sometimes reveals short-term trends against political authority such as #Erdoğandan Korkmuyorum (IAmNotAfraidOfErdogan), #Hükümet İstifa (GovernmentResign), #BilaleAnlatırGibi (AsIf ExplainingToBilal), #MilletNefesAlamıyor (PeopleCanNot Breath), #yönetemiyorsunuz (YouCanNotGovern), and #HepBirlikteArtıkYeterDiyoruz (WeSayEnoughAll

Together), but because it is not sustainable, it cannot reach the majority in a very short time and disappears. This indicates that with too much information and information over-consumption, "passive indifference" has become a cultural norm (Lovink, 2013, p. 6).

Today, Twitter is an area of struggle. This has gained more importance with the conscious and organized practices of political authority, especially after the Gezi Park protests. Although discourse against the political authority is said to have created an agenda via Twitter, in practice, contextual shifts with the performances of government trolls are remarkable. The populist language of communication does not make communicative action possible on Twitter. Thus, the reflexive counter-discourses developed by real people or trolls via Twitter sometimes detract the medium from participatory democracy since it resembles the AK Party's troll army's language of communication. Polarization, the main building block of the AK Party's populist politics—when troll and counter-troll posts are examined—has become clearer on Twitter. Nevertheless, the communication strategy of the group, which has been marginalized by the AK Party as the country's elite, is closer to the Habermasian politics of respect within the framework of the ethics of public discourse (Habermas, 1990). The majority of the 18,000 tweets categorized within five themes strives to "inform the public," "preserve the values of the republic," and "develop ideal discourses within the framework of the common good." This indicates that counter-discourses are striving to preserve the common good by trying to produce intersubjective consensus in the Boğaziçi protests. Similarly, this attitude of counter-discourses is similar to the Shils' (1991) relationship between civility and civil society because as a feature of civil society, respect takes care of not only the solicitude of the whole society, but also reveals a concern for the establishment of the common good (pp. 11–12). Hence, NGOs in Turkey need to develop a counter-action and communication strategy by being more active and organized. Since plurality and popularity are important on user-derived platforms such as Twitter, it is significant for artists, politicians, and other intellectuals to engage in discussions and for the public to achieve the common good. As observed in the research, discussions that real people do not participate in and do not support remain on the trends for a very short time. Besides, Twitter provides an opportunity to destroy the dominance of the tyranny of the majority created by the intricate relationship between traditional media ownership and political authority. It ensures the creation of public opinion by making protests and people visible, which are made invisible by the captured media. However, excessive information circulation and consumption causes us to ask the question of how organic the resulting public opinion is. How many people remember the Twitter trends—even the country's agenda—examined in the study? The establishment of a counter-hegemonic historical bloc stipulates an economic, ideological, and institutional organization, not just a discursive struggle. Counter-trolls' ability to form a new bloc against the hegemony of political authority depends on the existence of non-governmental organizations focused on communicative action, autonomous and strongly funded. As a result, the findings demonstrate that Twitter does not have sufficient potential for the establishment of the neo-Gramscian counter-hegemony, but there is also the potential for the communicative action taken by counter-trolls within the framework of peaceful actions and discourses to create public opinion (Konda, 2021) and mobilize the majority as in the Gezi Park protests.

Conflict of Interests

The author declares no conflict of interests.

References

Akgemici, E. (2019, February 7). Venezuela'da karşı-hegemonya: Yenilmeye mahkûm bir mücadele mi? [Counter-hegemony in Venezuela: A struggle doomed to defeat?]. *Birikim*. https://birikimdergisi.com/guncel/9341/venezuelada-karsi-hegemonya-yenilmeye-mahkum-bir-mucadele-mi

'Ak trol'lerin haritası çıkarıldı: Merkezde Erdoğan'ın danışmanı Varank var [Aktrolls' map has been drawn out: Erdogan's advisor Varank is at the center]. (2015, October 14). *Diken*. https://www.diken.com.tr/ak-trollerin-haritasi-cikarildi-merkezde-erdoganin-danismani-varank-var

Altuntaş, Ö. (2015, May 18). AKP sosyal medya merkezi'nde bir gün [A day at the social media center of AK Party]. *BBC News*. https://www.bbc.com/turkce/haberler/2015/05/150515_akp_sosyal_medya

Aytaç, S., Çarkoğlu, A., & Elçi, E. (2021). Partisanship, elite messages, and support for populism in power. *European Political Science Review*, *13*(1), 23–39.

Bajomi-Lázár, P. (2013). The party colonisation of the media: The case of Hungary. *East European Politics and Societies*, *27*(1), 69–89.

Baloğlu, U. (2019). Televizyon dizilerinde komedinin araçsallaşması: etik sınırların muğlaklaşarak toplumsal cinsiyetin yeniden üretimi [The instrumentalization of comedy in television series: Reproduction of gender by ambiguating of ethical boundaries]. In Z. B. Şahin (Ed.), *İletişim Etiği: Kavramlar, Olgular ve Tartışmalar* [Communication ethics: Concepts, facts and debates] (pp. 221–258). Literatürk.

Binark, F. M., Karataş, Ş., Çomu, T., & Koca, E. (2015). Türkiye'de Twitter'da trol kültürü [Troll culture on Twitter in Turkey]. *Toplum Ve Bilim*, *135*, 124–157.

Boğaziçi Dayanışması. [@boundayanisma]. (2021, February 12). *Direnişimizin 40. Günü, hocalarımızın basın açıklaması* [40th day of our resistance, press release of our teachers] [Tweet]. Twitter. https://twitter.

com/boundayanisma/status/1360178953206571012
Boğaziçi Direnişi. [@budirenisi]. (2021, February 1). *Eylemimizin barışçıl olduğunu tekrarlayıp taleplerimizi yineliyoruz!* [We reiterate that our protest is peaceful, we reiterate our demands!] [Tweet]. Twitter. https://twitter.com/budirenisi/status/1356284590768467970
Bovet, A., & Makse, H. A. (2019). Influence of fake news in Twitter during the 2016 US presidential election. *Nature Communications*, *10*(1), 1–14.
Brummette, J., DiStaso, M., Vafeiadis, M., & Messner, M. (2018). Read all about it: The politicization of "fake news" on Twitter. *Journalism & Mass Communication Quarterly*, *95*(2), 497–517.
Bulut, E., & Yörük, E. (2017). Mediatized populisms/digital populism: Trolls and political polarization of Twitter in Turkey. *International Journal of Communication*, *11*, 4093–4117.
Çam, A., & Yüksel, İ. Ş. (2015). Türkiye'de medyanın 2002 sonrası dönüşümü: Ekonomi politik bir yaklaşım [The transformation of the media in Turkey after 2002: A political economy approach]. In U. Aydın (Ed.), *Neoliberal Muhafazakâr Medya* [Neoliberal conservative media] (pp. 66–103). Ayrıntı.
Çiğdem, A. (2004). *Bir imkân olarak modernite: Weber ve Habermas* [Modernity as an opportunity: Weber and Habermas]. İletişim
Colleoni, E., Rozza, A., & Arvidsson, A. (2014). Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. *Journal of Communication*, *64*(2), 317–332.
conta cunte. [@cslecter]. (2020, March 31). *Eyyy Dünya liderleri..!* [Heeeey World leaders..!] [Tweet]. Twitter. https://twitter.com/cslecter/status/1245055660229832707
Değer, O. (2015). Entelektüel itaatsizlik ve politik mizah: Gezi direnişi [Intellectual disobedience and political humor: Gezi resistance]. *Mülkiye Dergisi*, *39*(2), 319–326.
Duran, R. (2015). Dümdüz….Sessiz….Hareketsiz….Ulvî…. Gibi….Medya nasıl kürtaj edilir? [Straight….Quiet…. Stable….Like….Divine….How to abort media?] In U. Aydın (Ed.), *Neoliberal Muhafazakâr Medya* [Neoliberal conservative media] (pp. 19–30). Ayrıntı.
Eemir Bereket. [@emirbereket]. (2021, February 2). *Sokakta, üniversitelerde, kamuda, mecliste TERÖRİST İSTEMİYORUZ* [WE DO NOT WANT TERRORISM on the street, in universities, in public, in parliament] [Tweet]. Twitter. https://twitter.com/emirbereket/status/1356660202624409602
Erem, O. (2020, July 2). Sosyal medya yasası: Türkiye'de bugüne kadar hangi siteler ve sosyal medya platformları yasaklandı? [Social media law: Which sites and social media platforms have been banned in Turkey so far?]. *BBC News*. https://www.bbc.com/turkce/haberler-turkiye-53259839
Fasce, A. (2020). The upsurge of irrationality: Pseudoscience, denialism, and post-truth. *Disputatio: Philosophical Research Bulletin*, *9*(13), 1–22.
Freedom of expression and the press in Turkey—281. (2021, February 13). *Expression Interrupted*. https://www.expressioninterrupted.com/freedom-of-expression-and-the-press-in-turkey-281
Fuchs, C. (2017). Donald Trump: A critical theory perspective on authoritarian capitalism. *TripleC: Communication, Capitalism & Critique*, *15*(1), 1–72.
Fukuyama, F. (1995). Democracy's future: The primacy of culture. *Journal of Democracy*, *6*(1), 7–14.
Ghergina, S., Mişcoiu, S., & Soare, S. (Eds.). (2013). *Contemporary populism: A contoversial concept and its diverse forms*. Cambridge Scholars.
Göle, N. (2000). *Melez desenler* [Hybrid patterns]. Metis.
Gramsci, A. (1971). *Selections from the prison notebooks of Antonio Gramsci*. International Publishers.
Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 US presidential election. *Science*, *363*(6425), 374–378.
Grossman, S., Akis, F. A., Alemdaroğlu, A., Goldstein, J. A., & Jonsson, K. (2020). *Political retweet rings and compromised accounts: A Twitter influence operation linked to the youth wing of Turkey's ruling party*. Stanford Internet Observatory.
Habermas, J. (1990). *Moral consciousness and communicative action*. MIT Press.
Habermas, J. (2001). *İletişimsel eylem kuramı* [The theory of communicative action]. Kabalcı.
Habermas, J. (2013). *Kamusallığın yapısal dönüşümü* [The structural transformation of the public sphere]. İletişim.
Heper, M. (2000). The Ottoman legacy and Turkish politics. *Journal of International Affairs*, *54*(1), 63–82.
Hermida, A. (2010). From TV to Twitter: How ambient news became ambient journalism. *Media/Culture Journal*, *13*(2). https://doi.org/10.5204/mcj.220
idilos bébé aka. youroti. [@kafayikirdim]. (2021, February 2). *Roses are red Melih is Bulu he stole an article and a rektörlük, too* [Tweet]. Twitter. https://twitter.com/kafayikirdim/status/1356703778959941636
Jha, C. K., & Kodila-Tedika, O. (2020). Does social media promote democracy? Some empirical evidence. *Journal of Policy Modeling*, *42*(2), 271–290.
Kaç Saat Oldu? [@KacSaatOlduTR]. (2021a, February 2). *Biz nereye bakacağımızı Atamız'dan öğrendik* [We learned where to look from our Father] [Tweet]. Twitter. https://twitter.com/kacsaatolduson/status/1356372772780466180
Kaç Saat Oldu? [@KacSaatOlduTR]. (2021b, February 1). *Niye susuyorsun BayKemal?* [Why are you silent, Mr. Kemal?] [Tweet]. Twitter. https://twitter.com/KacSaatOlduTR/status/1356003772925931523
Kapheros. [@sigaramcamel]. (2021, February 2). *Cehaletin kadrolaşmış hali, götünüzle deve izi yapsanız giremeyeceğiniz okullardaki çocukları aşağılık bir eziklikle kıskanıyorsunuz* [The staffed state of

ignorance, you are jealous of the children in schools that you would not be able to enroll even if you made a camel picture with your ass] [Tweet]. Twitter. https://twitter.com/sigaramcamel/status/1356654188571340805

Karatas, D., & Saka, E. (2017). Online political trolling in the context of post-Gezi social media in Turkey. *International Journal of Digital Television*, *8*(3), 383–401.

Karpat, K. H. (2010). *Türk demokrasi tarihi: Sosyal, ekonomik, kültürel temeller* [History of Turkish democracy: Social, economic, cultural foundations]. Timas.

Kaya, M. (2015). The modernization of Turkey as an example of non-Western modernization: Continuities, ruptures, and diversifications. *Turkish Studies*, *10*(2), 545–564.

Kemp, S. (2020). *Digital 2020: Global digital overview*. Global Digital Insights. https://datareportal.com/reports/digital-2020-global-digital-overview

Keyes, R. (2004). *The post-truth era: Dishonesty and deception in contemporary life*. Macmillan.

Konda. (2021). *Toplumun Boğaziçi Üniversitesi olaylarına bakışı* [The perspective of the society on Boğaziçi University protests]. https://konda.com.tr/tr/rapor/toplumun-bogazici-universitesi-olaylarina-bakisi

Kongar, E. (2001). Demokrasi kültürü sorunları [The problems of democracy culture]. *Emre Kongar'ın Resmi İnternet Sitesi*. https://www.kongar.org/makaleler/Demokrasi_Sorunlari.php

Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a social network or a news media? In M. Rappa & P. Jones (Eds.), *WWW '10: Proceedings of the 19th international conference on World wide web* (pp. 591–600). ACM. https://dl.acm.org/doi/10.1145/1772690.1772751

Laclau, E. (2005). *On populist reason*. Verso.

Le Bon, G. (1997). *Kitlelerin psikolojisi* [Psychology of crowds]. Hayat.

Levitin, D. J. (2017). *Weaponized lies: How to think critically in the post-truth era*. Penguin.

Lovink, G. (2013). After the social media hype: Dealing with information overload. *e-flux Journal*, *45*. https://www.e-flux.com/journal/45/60109/after-the-social-media-hype-dealing-with-information-overload

Mamo Dayı. [@AntepliMamato]. (2021, February 2). *-Boğaziçi "Sömürgeci" işgalinden kurtarıldı. -Askeriye "Fetö" işgalinden kurtarıldı* [The Boğaziçi was saved from the "Colonial" occupation. -The military was saved from the "Feto" invasion] [Tweet]. Twitter. https://twitter.com/AntepliMamato/status/1356516973308477441

Mardin, Ş. (1991). *Türk modernleşmesi* [Turkish modernization] (Vol. 4). İletişim.

Marginale. [@THEMARGİNALE]. (2021, February 1). *Alın işte demiştik. Dertleri ne rektör, ne de üniversite* [There you are, we said before. Their problems are neither rector nor university] [Tweet]. Twitter. https://twitter.com/THEMARGİNALE/status/1356338072426848257

McIntyre, L. (2018). *Post-truth*. MIT Press.

McLuhan, M. (1963). *The Gutenberg galaxy*. University of Toronto.

Media Ownership Monitor Turkey. (2019). *Media & owners database*. https://turkey.mom-rsf.org/en

Merrin, W. (2019). President troll: Trump, 4chan, and memetic warfare. In C. Happer, A. Hoskins, & W. Merrin (Eds.), *Trump's media war* (pp. 201–226). Macmillan.

MYazar212. [@MYazar212]. (2021, February 1*). Ne eğitimi yönetebiliyorsunuz..! Ne sağlığı yönetebiliyorsunuz..!* [Neither can you manage the education! Nor can you manage the health..!] [Tweet]. Twitter. https://twitter.com/MYazar212/status/1356311515666780160

Neuendorf, K. A. (2019). Content analysis and thematic analysis. In P. Brough (Ed.), *Research methods for applied psychologists: Design, analysis, and reporting* (pp. 211–223). Routledge.

Ongur, H. O. (2015). Identifying Ottomanisms: The discursive evolution of Ottoman pasts in the Turkish presents. *Middle Eastern Studies*, *51*(3), 416–432.

Özsoy, D. (2015). Tweeting political fear: Trolls in Turkey. *Journal of History School*, *12*(22), 535–552.

Pinker, S. (2018). *Enlightenment now: The case for reason, science, humanism, and progress*. Penguin.

Recuero, R., & Gruzd, A. (2019). Cascatas de fake news políticas: Um estudo de caso no Twitter. *Galáxia*, *41*, 31–47.

Reporters Without Borders. (2020). *World press freedom index of reporters without borders*. https://rsf.org/en/taxonomy/term/145

Schiffrin, A. (2018). Introduction to special issue on media capture. *Journalism*, *19*(8), 1033–1042.

Shils, E. (1991). The virtue of civil society. *Government and opposition*, *26*(1), 3–20.

SON LAİK BÜKÜCÜ. [@TheLaikYobaz]. (2021a, February 1). *Olimpos dağının korkak çocuklarına karşı, Hira Dağının cesur evlatlarıyız!* [Against the cowardly children of Mount Olympus, we are the brave sons of Mount Hira!] [Tweet]. Twitter. https://twitter.com/TheLaikYobaz/status/1355989064663658498

SON LAİK BÜKÜCÜ. [@TheLaikYobaz]. (2021b, February 2). *Kadıköy'de teröristler tarafından polise karşı barikat kuruluyor* [A barricade is erected by terrorists against the police in Kadıköy] [Tweet]. Twitter. https://twitter.com/TheLaikYobaz/status/1356661682609733634

The Journalists' Union of Turkey. (2021). *67 journalists jailed in Turkey*. https://tgs.org.tr/arrested-jailed-journalists-turkey

Turkish Statistical Institute. (2020). *The results of address based population registration system*. https://turkstatweb.tuik.gov.tr/PreTablo.do?alt_id=1059

Weedon, J., Nuland, W., & Stamos, A. (2017, April 27). *Information operations and Facebook*. [Press

Release]. https://www.mm.dk/wp-content/uploads/2017/05/facebook-and-information-operations-v1.pdf

Wodak, R. (2015). *The politics of fear: What right-wing populist discourses mean*. SAGE.

Yakut. [@Maesteroid]. (2021, February 2). *Boğaziçi öğrencileri üniversitelerde başörtüsü yasağını böyle protesto etmişti* [Boğaziçi students protested the headscarf ban in universities] [Tweet]. Twitter. https://twitter.com/Maesteroid/status/1356530483807588352

Zenibya. [@bensuse_]. (2021, February 2). *BU VAHŞET GÖRÜLSÜN. KİMSE GÖZÜNÜ KAPAMASIN. KİMSE SUSKUN KALMASIN* [THIS VIOLENCE SHOULD BE NOTICED. NO ONE SHOULD IGNORE IT. NO ONE SHOULD BE SILENT] [Tweet]. Twitter. https://twitter.com/bensuse_/status/1356642979985948680

About the Author

Uğur Baloğlu is an assistant professor in the Faculty of Applied Sciences at Istanbul Gelişim University. He completed his MA in 2012 in the field of communication design at Istanbul Kültür University. He completed his PhD in radio, tv, and cinema at Istanbul University in 2017. He is the co-editor of *Transcultural Images in Hollywood Cinema: Debates on Migration, Identity, and Finance* (2021). His research interests include media, cultural studies, and communication research.

Media and Communication (ISSN: 2183–2439)
2021, Volume 9, Issue 4, Pages 52–61
https://doi.org/10.17645/mac.v9i4.4225

Article

Media Control in the Digital Politics of Indonesia

Masduki

Department of Communication, Universitas Islam Indonesia, Indonesia; E-Mail: masduki@uii.ac.id

Submitted: 26 February 2021 | Accepted: 11 June 2021 | Published: 21 October 2021

Abstract
In transitional democratic countries with significant digital media user bases, the "authoritarian turn in digital media" has resulted in new forms of media control designed to counter critical media exposure. This article investigates the ongoing digital pressures experienced by Indonesian media organizations and investigative journalists by the partisan supporters of the country's new authoritarian political leaders. This article provides a critical review of the forms of media control that have emerged in Indonesia within the past five years (2015–2020), giving special attention to the doxing allegedly faced by several news media and journalistic projects: IndonesiaLeaks; *Tempo* magazine; and WatchDoc. Applying qualitative methods (observation, semi-structured interviews, review of documents), this study finds that the rise of non-state and societal control over critical media leads to self-censorship amongst media and journalists. This study shows that online trolls, doxing, and hyper-partisan news outlets are used as new forms of media control. Control is also exerted by paid-social media buzzers, whose online identity is established by their use of digital and social media platforms to manipulate information and counter critical news regarding incumbent and oppositional political leaders. This article contributes to the academic debate on the intended forms of media control in digital politics of transitional democracies.

Keywords
digital politics; doxing; Indonesia; journalism; media buzzers; media control

Issue
This article is part of the issue "Media Control Revisited: Challenges, Bottom-Up Resistance and Agency in the Digital Age" edited by Olga Dovbysh (University of Helsinki, Finland) and Esther Somfalvy (Research Centre for East European Studies at the University of Bremen, Germany).

1. Introduction

In linking the rapid growth of social media platforms with democracy and press freedom, scholars have fallen into two categories. Some academics portray social media as a key driver of an emerging media ecosystem that circles around public participation and democratic accountability (Jenkins et al., 2012; Paterson, 2019). Social media platforms have been hailed as potential saviors of news production, allowing journalists to find new informants and data sources, thereby engaging directly with their preferences and interests (Hermida, 2011; Johansson, 2016). Internet platforms may relieve the press from its reporting obligations, leaving the press free to focus on investigative journalism projects (Anderson et al., 2012).

Other scholars, however, argue that these platforms —rather than allowing users to contribute information and observations to news production—allow anti-democratic groups to control and manipulate news (Poell & van Dijck, 2014). Rather than using social media as a tool of deliberative communication (Kangei, et al., 2018), paid-social media users (popularly named buzzers) use free access online platforms to control news and use hate speech to degrade quality journalism.

With those background in mind, this article interrogates the new mechanisms of media control through which social media activity affects the news process and threatens media organizations. This article seeks to scrutinize the practices, motives, and actors of digital threats used to control the media in Indonesian digital politics within the past five years (2015–2020). Special attention will be given to the doxing allegedly faced by several media organizations and journalistic projects: IndonesiaLeaks, a joint collaborative platform

of several media organizations to investigate corruption among Indonesia's police elite; *Tempo*, a leading investigative magazine during the revision of the Corruption Eradication Commission Law and the Job Creation Law; and WatchDoc, a documentary production house that explores the links between energy companies and Indonesia's top politicians. This article argues that, rather than enhancing journalistic freedom and media autonomy, the rise of social media has intensified the political pressures experienced by journalists.

Bradshaw and Howard (2019) found organized social media manipulation during Indonesia's 2014 and 2019 Presidential elections, but failed to observe its impact on media independence. Other studies (e.g., Irawanto, 2019; Johansson, 2016; Lim, 2017, 2018; Saraswati, 2018), have used similar approaches to explore the rise of social media buzzers in online political contestations with no observation on their implication to media freedom. Advancing the previous studies, this article will show that both the growing number of buzzers and the practice of media manipulation are new tools for controlling critical media. Doxing and surveillance are particularly used as new forms of media control in Indonesia during 2015–2020. Thus, this study offers an example of a transitional democratic country with a large digital media userbase and its efforts to minimize critical media exposure. This article then aims to re-consider and re-assemble the notion of media control, thereby using Indonesia as a case study to develop a new perspective for understanding media control in the digital age.

This article is ordered as follows. Following this introductory section, it revisits models of media control in the digital age within the context of new authoritarian politics. It also surveys the rise of organized social media manipulation, online trolls, etc. The third section of this article describes the method used in this study, while the fourth and fifth sections offer empirical findings and discussion, with a focus on the doxing allegedly faced by journalists/activists within IndonesiaLeaks, *Tempo* magazine, and WatchDoc. The sixth section provides this article's conclusions.

2. Media Control Revisited

This section elaborates on threats faced by critical media and journalists since the arrival of digital communication technology, viewing them as a form of media control. The term control is generally defined as the power to influence or direct the behavior of a person or agency. Individuals use various ways to exercise this power and negotiate the political interactions within organizations (Bicer, 2020). Control covers various formal and informal arrangements, both centralized and decentralized (Jingrong, 2010). The media, meanwhile, is an institution with the central aim of producing and distributing knowledge in the broadest sense of the word (McQuail, 2010), with journalists as its key actors. In the media sector, control can be exerted through policy, money, and practices, through which political authorities can exercise control over publication licenses, newsrooms, or journalists.

In both democratic and autocratic politics, media control is a key means through which authorities limit the public's access to critical information, thereby securing their political power. In African countries such as Uganda, media and journalists face several threats which include: state intimidation, arrest of those considered critical to the state, and denial of access to public information (Walulya & Nassanga, 2020). According to Google researchers, 21 of the world's top 25 news organizations have been targeted by hacking attacks, likely by state-sponsored actors (Wagstaff, 2014). Danger is faced not only by those who publish online, but also actors whose journalistic activities interface with digital platform, whether by using computers to process information, using the internet for news gathering and/or research, or simply relying on email for external contact. At an individual media worker, from 1992 to 2021, the Committee to Protect Journalists (CPJ) found 1398 journalists killed including who worked for online media (CPJ, 2021). The CPJ as well as the International Freedom of Expression Exchange have actively documented online attacks against journalists and media organizations, often committed by actors seeking to further sociopolitical goals.

Computational propaganda has increased the media and journalists' risk of digital threats. Threats commonly faced by media outlets include doxing and surveillance, software and hardware exploits, phishing attacks, fake domain attacks, intimidation, harassment, forced exposure of online networks and disinformation, confiscation of journalistic products, and data storage and mining (Henrichsen et al., 2015; Posetti, 2018). Doxing refers to the search for and publication of private or identifying data (about a particular individual) on the internet, typically with malicious intent (Douglas, 2016). Surveillance—the monitoring, interception, and retention of information—is one way that social media buzzers monitor information and the media. Regan (2012) argues that surveillance is often carried out on the grounds of determining power, and often exists without clear policies regarding this information.

Posetti et al. (2020) further found that artificial intelligence technology is being leveraged to create "deepfake" videos and other content designed to discredit targets, including journalists, and especially female reporters. In the case of investigative reporting, when the people being investigated are influential individuals in government, critical media organizations and journalists become the targets or indirect victims (Cottle, 2017). The surveillance technologies are globally diverse, and can include location tracking, facial recognition and monitoring, and bulk interception methods for voice, email, fax, and satellite phone calls to media.

Another common form of media control is the manipulation of information to undermine credible journalism and reliable information. Rumata and Sastrosubroto

(2018) found that manipulated information is increasingly presented as valid information. Selective information is reproduced and reframed as fact, then redistributed within certain groups via social media. This results in the deliberate targeting of media companies and media professionals, along with their sources, who seek to verify or share critical news and commentaries. More broadly, journalists and media organization have been targeted by acts of "trolling"—deliberate attempts to "misinform or endanger" by sharing information designed to distract their news sources. Journalists might be targeted to trick them into sharing inaccurate data, which feeds a false analysis of the facts or, when it is revealed to be fake, weakens the integrity of their news media.

In the "analog political" era, media control was realized by state officials through direct state censorship, advertising and strict media policies. Meanwhile, in the digital age, media control travels beyond its traditional models. When "traditional" and "new" media technologies have emerged simultaneously in many developing democracies, forms of media control do not replace one another, but combine and compete (Atal, 2017). The actors are no longer state authorities but a network of non-state and paid social media buzzers and/or partisan digital influencers who are seeking both political and economic benefits. They may be considered as a "criminal network." The criminal networks are politically prepared to organize digital violence to control information and investigations that threaten the interests of incumbent political leaders. Unaccountable internet corps and paid influencers play a central role in producing manipulated content to the public. For instance, online abuse and hate speech against journalists, including threats such as phishing, malware, and cyber espionage are on the rise, and have been disseminated via social media and mobile devices (Henrichsen et al., 2015). At the same time, traditional threats continue to occur offline and have found new ways to cause harm and create the hostile environment faced by journalists.

Political motivated buzzers (Felicia & Loisa, 2019) have become key players in challenging the work of media professionals. Although there is no formal definition, in the context of digital communication, the term buzzer is often used to refer to those with the capacity to influence others via social media, enliven online conversations through their tweets, voice their interests, and get paid for their postings. Buzzers were initially actors seeking to cultivate views or marketers seeking to sell products (Arianto, 2020; Saraswati, 2018; Sugiono, 2020). Yet, they have since become identified with negative promotional strategies used to spread political propaganda.

Buzzers are born and benefit from the social media userbase. They can be seen as autonomous and/or state-supported political actors, incorporated in political campaign strategies to increase the electability and popularity of specific political figures or parties. Bradshaw and Howard (2019) describe that buzzers tend to use automated and human labor to manage fake accounts, through which they convey support for candidates and attack their opponents, thereby polarizing (dividing) society. Such accounts may also be used to disseminate disinformation. Jati (2017) describes buzzers as working to produce *kul-twit* (twitter lectures) or mini-stories using academic and technocratic language, distributing messages using anonymous accounts, and "testing the water" of politics. Their discourse is only temporary, being used to gauge the actions and responses of middle-class social media users.

It can be concluded that digital politics present new forms of media control. Seeking to ban or at least delegitimize journalistic works (critical news to political authorities), social media buzzers employ threats to media and their journalists. In this sense, control over media acts to undermine their role in advancing the interests of the public. Threats commonly faced by media outlets include doxing and surveillance, software and hardware exploits, fake domain, partisan websites, and personal intimidation. Online threats may be made individually, or combined with offline ones in the interest of chilling critical views and quashing media freedom (Ruffini, 2018). At the same time, physical abuse—potentially fueled by online incitement and designed to suppress analytical reporting—continues (Posetti, 2018).

Such attacks are motivated and influenced by a variety of factors and interests, including the politics, social, and economic contexts where information controls are applied. They are also influenced by the available communications infrastructure, such as the number of internet service providers, and the overall level of internet penetration and growth. As for motive, such attacks desire to distract journalists and media outlets from their investigations by prompting fruitless lines of inquiry that stymie reporting efforts and, ultimately, have a chilling effect on truth-seeking. Examples of this style of misdirection include the struggled reframing of claims about the size of the crowd at Donald Trump's inauguration in 2017 as alternative facts (Smith, 2017). The fabrication of the event designed to trick journalists and US citizens, along with structured social media campaigns aimed at mimicking public response.

3. Method of Study

This article is driven by two research questions: What forms of media control have been used within Indonesia's growing digital political propaganda (2015–2020), and how has digital violence been used against media organizations and journalists operating in Indonesia to control their news coverage? Using Indonesia's digital politics as a backdrop, this article explores new patterns of media control based on the findings of a qualitative study conducted between 2019 and 2020.

Using a qualitative descriptive analysis (Yin, 2014) with a critical perspective, this article examines as much data as possible to describe and explain the diverse

forms of digitized media control practiced in Indonesia. The qualitative method was used to recognize cases of threats, analyze patterns of media control, and understand the extent to which digital threats are organized by cyber attackers. Analyzed data were collected through observations of selected news websites and social media platforms, semi-structured interviews, and reviews of pertinent documents.

Recognizing the huge number of Indonesian media, as well as their broad experiences with digital threats, this article selected three cases: IndonesiaLeaks; *Tempo* magazine; and WatchDoc. IndonesiaLeaks is an independent whistleblower platform founded by several media outlets to investigate corruption among Indonesia's police elite. *Tempo* is a leading news magazine and news website that regularly conducts investigative journalism. Finally, WatchDoc is an independent production house that produces critical documentaries, such as *Sexy Killers* (2019)—a news documentary exploring the links between energy companies and Indonesia's top politicians that was viewed by 20 million online viewers. These cases were selected to represent three types of journalistic works: a collaborative platform of several news media (IndonesiaLeaks); a recognized news media organization (*Tempo*, established in 1971); and an individual and independent news production house (WatchDoc).

Following Creswell (2014), this study was conducted in several stages. First, the author collected published documents on acts of digital intimidation committed against media between 2015 and 2020. We assessed official reports from Indonesian media authorities (e.g., Press Council [*Dewan Pers*], Ministry of Communication and Information Technology), then compared them by reading the annual reports of press, broadcasting, and internet freedom advocacy agencies in Indonesia (e.g., Alliance of Independent Journalists, and SAFEnet). During this stage, we reviewed several aspects of digital violence: number of cases, form, technology used, actor, target, and motive. A longitudinal report by the World of Journalism Study (Muchtar & Masduki, 2016) provided a broad picture of Indonesian journalistic culture and the challenges faced.

Second, complementing this desk review, between January and July 2020 the researcher conducted semi-structured interviews with 10 media professionals who had experienced online intimidation. Particular focus was given to news media executives and individual journalists with *Tempo* magazine and website, IndonesiaLeaks, and WatchDoc. Informants included; Abdul Manan, a senior journalist to *Tempo*, chairman of IndonesiaLeaks, and president of the Alliance of Independent Journalists (AJI); Shinta Maharani, the chair of AJI's Yogyakarta branch; Dandhy Dwi Laksono, the owner of WatchDoc; and Heru Margianto, the managing editor of Kompas.com. Interviews were designed with open-ended questions that allowed informants to express their views regarding the rapid rise of social media buzzers/influencers during Indonesian political events (2014–2019), current cases of digital violence against critical media outlets, their individual experiences with threats, and the correlation between said cases and buzzers.

Third, further materials were collected by observing several famous social media platforms, with a particular focus on two types: the social media platforms of individuals who are publicly identified as buzzers, and the social media platforms/websites of anonymous owners and/or operators. Specific attention was paid to their ownership of the platforms, legal position, and conversations on critical news media. In addition, this study observed the official websites of *Tempo*, Tirto.id, Liputan6.com, and Kumparan, all of which had experienced hacking and information manipulation. Special observation was paid to the news items published.

Finally, conclusions were drawn based on the collected materials and analytical concepts. The typologies of digital intimidation against media actors proposed by media researchers (e.g., Nadzir, 2020; Posetti, et al., 2020) have been used to identify examples of media control, while the views of political scientists (such as Power, 2018) have been used to place the cases in the context of Indonesian politics.

4. Findings

This section consists of two features. First, it provides a short description of Indonesia's politics and media system, thereby elucidating the backdrop of this study. Second, to answer its questions on the forms of media control and practice of digital violence against journalists and media organizations in Indonesia, this section explores the use of digital violence as a means of media control in the country's digital politics (2015–2020). The analysis focuses on three cases: *Tempo* group (both the magazine and the website), IndonesiaLeaks, and WatchDoc.

By May 1998, following the end of Suharto's authoritarian regime, Indonesia had embarked on a mission to adopt a liberal democratic political system. Ultimately, however, the country transitioned from strong direct state control to a more complicated form of political control (Tapsell, 2015; Warburton & Aspinall, 2019). Today, Indonesian control is marked by oligarchic practices and the rise of digital political contestations between state and non-state actors.

In Indonesia's first ten years of political reorganization (1998–2008), various reforms were implemented in key sectors—including the previously autocratic media system. This strengthened freedom of expression and public participation in the media. Several strategic policies were enacted, including laws on press (1999), broadcasting (2002), and access to information (2008). Combined, these introduced liberal pluralism in media ownership, content production, and deliveries.

At the same time, the Indonesian government welcomed democratic media principles through a series

of pro-democratic media laws, including Law No. 5 of 1999 on competition, Law No. 40 of 1999 on the press, and Law No. 32 of 2002 on broadcasting (Ministry of Communication and Information Technology of Indonesia [MCIT], 2020a, 2020b). At the time of legislation (1998–2002), political authorities shared the vision that several media policies were necessary to guarantee media independence. However, when the new authorities consolidated during the second decade of political reform, promoting media freedom remained a "big job," as it was challenged by the state-controlled culture that persevered in the new political climate, the monopolistic media ownership, and the rapid rise of paid political cyber-troops.

On the regulatory side, there were contradictions in media policy. For instance, the Press Law offers protections to journalists, but the Information and Electronic Transactions Law (Law No. 19 of 2016, also known as the ITE Law, MCIT, 2020c) contains articles that threaten journalists with imprisonment, and indeed the law has been used for direct attacks against media professionals. The ITE law limits online journalistic practices by threatening journalists with up to six years' imprisonment or fines of up to one billion IDR ($106,000) for online defamation. For instance, SAFEnet, a digital rights defender throughout Southeast Asia, notes that at least 14 charges were levied against media organizations and journalists between 2008 and 2020 (SAFEnet, 2021). Further, SAFEnet (2021) finds the revision of the ITE law, discussed in the Indonesian policymakers between 2020–2021, could also provide the ruling government with a major tool to control news media and promote violence against human rights activists.

Indonesia is home to more than 175 million social media users that spend an average of 4.5 hours per day connected to internet. These fantastic numbers have proven attractive to politicians, which is reflected in the emergence of buzzers. Although buzzers are commonly defined as celebrities with at least 2,000 followers, the case in Indonesia is quite different: According to one Reuters' article, Indonesia's buzzers are not only celebrities, but also "ordinary people" or community members (Tapsell, 2019). As noted by Tapsell (2019), both Prabowo and Joko Widodo (Jokowi) clearly had "social media buzzer" teams running to shape digital discourse while concurrently countering and producing hoaxes and "black campaign" material. The Indonesia Corruption Watch has discovered that, since 2017, the Jokowi's administration has spent Rp 90,45 billion to fund influencers. For instance, as of late 2020, the government has yet to admit that it paid influencers to endorse the job creation bill (*UU Cipta Kerja*). This bill was criticized for prioritizing business interests at the expense of facilitating exploitative labor and ecological degradation.

Freedom House (2020a) reported that both Jokowi and Prabowo hired online campaign strategists to mobilize paid commenters and automated accounts to spread political advertising ahead of the 2019 election. The agency claims that one buzzer operated approximately 250 fake social media accounts on platforms such as Facebook, YouTube, and Twitter, and these and similar spam accounts amplified hashtags to benefit specific presidential candidates and control public interest media and journalism.

Freedom House (2020b) reports that, overall, press freedom in Indonesia has gradually declined, with the country's rank decreasing from 57 in 2002 to 124 in 2018 (from "free" to "partly free"); the country was again ranked "partly free" in the 2019 survey, which may be attributed to the rise of partisan websites, disinformation, and doxing against media professionals. Similarly, Reporters Without Borders (2018) ranked Indonesia 119th out of 179 countries, a significant drop from its 2002 peak (57th of 139 countries). As such, even as internet penetration has increased steadily, Indonesian media actors and journalists face new forms of control.

The following discussion exposes three popular cases of media control that have contributed to Indonesia's decreased press freedom, thereby showcasing how digital communication is used in efforts to control media and critical information. Investigative journalistic works are the main targets of harassment. *Tempo* magazine—Indonesia's leading investigative news agency—is the most targeted media for cyber intimidation, followed by IndonesiaLeaks and WatchDoc. Recognizing this condition, this article's discussion will begin with a discussion of the *Tempo* magazine case.

Three types of digital attacks—hacking, doxing, and surveillance—have been organized by digital attackers to control *Tempo* magazine. Each received public attention and invited public protests. *Tempo*, established in 1971, had previously been forced to shut down in 1994 amid claims by the Suharto power that the media threatened national stability. Publication of the magazine could not be resumed until 1999. Between 2015 and 2020, the magazine faced much online abuses over its investigative journalism on acting political powers. For instance, in mid-2020, the magazine's homepage—https://majalah.tempo.co—was hacked and changed with accusations that it was promulgating fake news (referring to its previous investigation of the network of buzzers supporting President Jokowi's Job Creation Law). Action was also taken by buzzers in response to *Tempo*'s political news allegation that the president was attempting to establish a political dynasty after his son Gibran Rakabuming Raka and son-in-law Bobby Nasution contested (and won) mayoral races in Solo and Medan (*Tempo*, 2020). On an almost weekly basis, *Tempo* has been threatened for its criticism of the Jokowi's political administration. The cover of the magazine's September 14, 2019, issue attracted public outrage for depicting Jokowi alongside the silhouette of Pinocchio as part of its weekly investigative report entitled "*Janji Tinggal Janji*" ("promises remain promises"; *Tempo*, 2020).

Since 2014 to 2020, the magazine's webpage Tempo.co has faced doxing and surveillance. Several attacks have been found using various hashtags, such as #TempoAsu. One group—calling itself Zone Injector—gained control of the homepage and replaced it with the phrase "we warned you, but you did not respond to our good intentions"; the Tempo.co website was down for five minutes (Pebrianto, 2020). Similarly, Tirto.id—a leading data journalism website—found that its site had been infiltrated in 2020. The buzzers replaced several news articles, including those critical of the Jokowi authority's handling of Covid-19 pandemic (Salam, 2020).

At the same time, doxing and surveillance have also been used against the individuals (sources) who provide insight to *Tempo* and other critical media. Take, for example, the experiences of anti-corruption activists Oce Madril and Rimawan in late 2019, and epidemiologist Pandu Riono. Rahayu et al. (2019) describe this as indicative of a pattern to silence public criticism. Ravio Patra, an independent researcher who had previously been a vocal critic of the Jokowi's young staff, was suddenly imprisoned and reportedly charged with spreading offensive messages on his WhatsApp account. His phone had been hacked, and the messages posted by the hackers (Nadzir, 2020). This action can be seen as a tactic to control public information and reduce quality of journalism provided by *Tempo* (Couper & Andriyanto, 2021).

The second popular case of media control is a digital violence to IndonesiaLeaks, a joint collaborative digital platform for investigative journalism. As a tactic to counter the coalition's investigative reporting, its chairman Abdul Manan (the president of the Alliance of Independent Journalist) was reported to the Jakarta Police on October 23, 2018, for a criminal case and to the South Jakarta District Court for a civil lawsuit on October 24, 2018. These reports were filed after the IndonesiaLeaks published an investigative news piece in December 2017. Called *The Red Book Scandal*, this report exposed evidence showing that significant amounts of money had been transferred from Indonesia's Corruption Eradication Commission investigators (former senior members of the Police Corps) to elite police officers (Global Investigative Journalism Network, 2018). The story also claimed that bribe money had flowed to General Tito Karnavian, the National Chief of Police.

Following IndonesiaLeaks, five of IndonesiaLeaks' nine media partners—*Tempo*, Kantor Berita Radio 68H (KBR), Suara.com, Independen.id, and Jaring.id—faced online bullies after they published the story across several platforms. The report as well as the media was quickly targeted by social media buzzers questioning the report, with hashtags #IndonesiaLeakshoax and #petisihoax. KBR's website was hit by a denial-of-service attack, leaving it inaccessible for a few hours.

The digital threats are used to control not only conventional media outlets, but also online-only and small-scale critical media houses, and even freelance journalists-cum-activists. For instance, digital risks to WatchDoc, an independent and Jakarta based-production house. One of the most controversial cases was the surveillance, doxing, and organized police reporting of the WatchDoc founder Dandhy Dwi Laksono for his series of critical journalistic works. Laksono is the producer of *Sexy Killers* (2019), a critical documentary that explores the links between energy companies and Indonesia's top politicians. During its production and public screenings, the WatchDoc platform and its journalists-cum-activists faced several internet trolls of their online posts and physical activities (Prabowo, 2019).

For instance, in the interest to counter his criticism, Laksono was charged with spreading hate speech by the Jakarta police on September 26, 2019, after posting about conflicts in two biggest cites of Papua province—Jayapura and Wamena—on his Twitter account (Dipa, 2019). Laksono was accused of violating Article 28 and 45 of the Electronic Information and Transactions Law and spreading information to fire hatred based on race (Dipa, 2019). On 23 September 2019, he had written about the Papua conflicts including a photo of students who had allegedly been shot during the incident. Laksono's arrest came three weeks after human rights activist Veronica Koman was named a suspect by police after she posted on Twitter account in support of the protesting Papuans, prompting rights groups to condemn the police action (Lamb, 2019).

5. Discussion

Media control, it can be seen, is exercised in the digital politics of Indonesia. Between 2015 and 2020, doxing, threats, and surveillance of critical media have all been commonly used for control. In the Suharto's authoritarian era (1960s–1990s), media control had been practiced through direct phone calls to media newsrooms and blackmail, as experienced by *Kompas* daily in 1965 and 1978, *Tempo* magazine in 1982 and 1994, and *Editor* magazine and *Detik* tabloid in 1994. Significant changes are thus occurring in media control after 1998, not only in the technology, packaging, or tactics used, but also in the actors involved. However, the motivation remains similar: to manage the credibility of the ruling authorities (Ruffini, 2018).

From the above data, we can say that *Tempo* magazine, IndonesiaLeaks, and WatchDoc offer examples of how media operating in Indonesia have been attacked by digital violence, thereby resulting in control of their news services. Unlike in 1982 and 1994, when the perpetrators of raids were clearly identified, in the digital era hackers' identities are not known. Indeed, in several recent cases, attacks have targeted not only media institutions but also groups of journalists (Parikesit, 2020).

This study finds that control of media in the digital politics is not centralized amongst state administrators. Control is exerted primarily by non-state and

social media-based buzzers—those whose online identity is established by their manipulation of information on various digital and social media platforms to counter critical news regarding incumbent and oppositional political leaders. They are also considered as paid influencers, which are formally or informally recruited by key political leaders as well as the ruling political administration (Wahidin & Ridwan, 2020). In the 2014 and 2019 presidential elections, Jokowi and his rival Prabowo Subianto both used social media buzzers or cyber-troops to promote their political campaigns and led many controversial attacks against media and journalists.

At the macro level, the increased control of media and public critics marked the authoritarian turn of Jokowi's politics (Power, 2018). The author sees that Jokowi's political power has taken the authoritarian turn ahead of the 2019 elections through manipulation of powerful law enforcement and security institutions for narrow, partisan purposes, and his political cyber-troop's concerted efforts to block critics of oppositional leaders and human right activists. This leads to the increasingly disempowerment of political opposition through a practice of digital repression that undermines freedom of online political communication and reduces political culpability.

Nadzir (2020) has noticed that, historically, digital platforms have always been integral to President Jokowi's political campaigns, at least since his run in the Jakarta gubernatorial election (2012). He believed that online platforms were crucial political tools. In this sense, his regime's recruitment of social media influencers is not surprising. Nadzir (2020) further finds that, by continuing to fund the campaign after winning the 2019 election, the Jokowi's government risks transferring the digital attacks to media professionals and media organization into day-to-day politics.

Parikesit (2020) notes that these attacks are intended to interfere with the media's work and potentially damage media actors' relationship with their sources or interviewees. It is broadly clear that such action could potentially interfere with freedom of expression, especially within the context of digital rights—i.e., (1) the right to access, (2) the right to expression, and (3) the right to feel safe.

To counter digital attacks, media houses, journalist associations, and the Press Council of Indonesia have organized various actions, including proactively exposing these attacks, filing official reports with police, and exposing attackers, thereby protecting their sources from further online victimization. AJI has regularly monitored the practice of doxing against journalists, and noticed that such actions usually result in persecution. To stop the trend, AJI and other non-profit press freedom agencies joined the Anti-Persecution Coalition in 2017 (Putra et al., 2018). The coalition has formed a crisis center to protect as well as provide legal assistance to victims of persecution and harassment. Meanwhile, responding to a series of digital attacks to control their public service, *Tempo* has stood by its journalistic principles and avoided criticism based on hatred or political motives. Furthermore, several media agencies have continued to respect the right of reply and covered both sides of stories.

6. Conclusion

This article has identified several forms of media control using Indonesia's digital platforms, including doxing, online trolling, surveillance, and information manipulation. It confirms that control of media in the digital era differs significantly from control in the analog era. Actions are organized by non-state actors, individuals, or—commonly—by social media buzzers closely involved with autocratic political leaders. This has signaled an authoritarian turn in Jokowi's politics, yet, it was not severe enough to mark the regime as transitioning democracy.

Through desk reviews and interviews with media executives and journalists, it was found that both media organizations and their informants are threatened by "digital disturbances." It further indicates that, even though journalism has become increasingly digital, Indonesia has become no safer for those expressing critical opinions. Journalists and media actors can reach their audiences more quickly, but threats—both new and old—await them: doxing, police reports, and surveillance. Through digital attacks, media outlets risk retaliation from non-state groups—particularly buzzers motivated to foster media distrust and escalate political instability.

The findings of this article contribute to the academic debate on the forms of media control exercised in digitalized political culture, mainly in Indonesia. However, this article has only addressed three case studies of media control, excluding other news media and journalistic works within the context of Indonesia's digital journalism. More importantly, this article has concentrated specifically on Indonesia, and thus extended investigation of media control in other transitional democratic countries is needed to compare the forms of media control used in digital politics. Given recent tendencies for media liberalization and the rise of digital/social media use, it remains challenging to study the media freedom and media control practiced by state authorities against non-state and digital-based actors in post-repressive societies.

Overall, this study showed how the digital threats control journalists' activism and critical news media. To agree with Nadzir (2020), whether through doxing, trolling, deactivating personal accounts, or removing articles from news sites, online attacks threaten media autonomy and critical voices in society. When critical voices are bullied into silence, raising public criticism involves considerable risk.

Conflict of Interests

The author declares no conflict of interests.

References

Anderson, C. W., Bell, E., & Shirky, C. (2012). *Post-industrial journalism: Adapting to the present—A report*. Columbia Journalism School.

Arianto, B. (2020). Salah kaprah ihwal buzzer: Analisis percakapan warganet di media sosial [An analysis of citizen conversation in social media]. *Jurnal Ilmiah Ilmu Pemerintahan*, *5*(1), 1–20. https://doi.org/10.14710/jiip.v5i1.7287

Atal, M. (2017). Competing forms of media capture in developing democracies. In A. Schiffrin (Eds.), *In the service of power: Media capture and the threat to democracy* (pp. 19–31). CIMA.

Bicer, C. (2020). The power and politics in organizations. In U. Ucan (Eds.), *Discussions between economic agents-socio economic studies* (pp. 221–245). Iksad Publications.

Bradshaw, S., & Howard, P. (2019). *The global disinformation order: 2019 global inventory of organized social media manipulation*. Oxford Internet Institute.

Committee for Protect Journalist. (2021). 1398 journalists killed between 1992 and 2021. https://cpj.org/data/killed/?status=Killed&motiveConfirmed

Cottle, S. (2017). Journalist killings and the responsibility to report. In U. Carlsson & R. Pöyhtäri (Eds.), *The assault on journalism: Building knowledge to protect freedom of expression* (pp. 21–33). Nordicom.

Couper, E., & Andriyanto, H. (2021, January 17). How Tempo outrages Jokowi supporters. *Jakarta Globe*. https://jakartaglobe.id/news/how-tempo-outrages-jokowi-supporters

Creswell, J. (2014). *Research design: Qualitative, quantitative and mixed methods approaches* (4th ed.). SAGE.

Dipa, A. (2019, September 27). Filmmaker Dandhy Laksono named 'hate speech' suspect for tweeting about clashes in Papua. *Jakarta Post*. https://www.thejakartapost.com/news/2019/09/27/filmmaker-dandhy-laksono-named-hate-speech-suspect-for-tweeting-about-clashes-in-papua.html

Douglas, D. (2016). Doxing: A conceptual analysis. *Ethics and Information Technology*, *2016*(18), 199–210. https://doi.org/10.1007/s10676-016-9406-0

Felicia & Loisa, R. (2019). Actor network and cohort cultures in the business of political buzzer. In A. P. Irawan (Ed.), *Proceedings of the Tarumanagara International Conference on the applications of social sciences and humanities (TICASH 2019)* (pp. 309–315). Atlantis Press.

Freedom House. (2020a). *Freedom on the net 2020: Indonesia*. https://freedomhouse.org/country/indonesia/freedom-net/2020

Freedom House. (2020b). *Freedom in the world 2020: Indonesia*. https://freedomhouse.org/country/indonesia/freedom-world/2020

Global Investigative Journalism Network. (2018). *IndonesiaLeaks: Officials attack first investigative report from whistleblower platform*. https://gijn.org/2018/10/25/indonesialeaks-officials-attack-first-investigative-report-from-whistleblower-platform

Henrichsen, J., Betz, M., & Lisosky, J. (2015). *Building digital safety for journalism: A survey of selected issues*. UNESCO.

Hermida, A. (2011). Mechanisms of participation: How audience options shape the conversation. In J. Singer, D. Domingo, A. Heinonen, A. Hermida, S. Paulussen, T. Quandt, Z. Reich, & M. Vujnovic (Eds.). *Participatory journalism: Guarding open gates at online newspapers* (pp. 13–33). Wiley.

Irawanto, B. (2019). Making it personal: The campaign battle on social media in Indonesia's 2019 Presidential election. *ISEAS Yusof Ishak Institute*, *2019*(28), 1–11. https://www.iseas.edu.sg/images/pdf/ISEAS_Perspective_2019_28.pdf?

Jati, W. (2017). Aktivisme kelas menengah berbasis media sosial: Munculnya relawan dalam Pemilu 2014 [Middle class activism in social media: The emergence of volunteerism in the 2014-presidential election]. *Jurnal Ilmu Sosial Dan Ilmu Politik*, *20*(2), 147–162. https://doi.org/10.22146/jsp.24795

Jenkins, H., Ford, S., & Green, J. (2012). *Spreadable media: Creating value and meaning in a networked culture*. NYU Press.

Jingrong, T. (2010). The crisis of the centralized media control theory: How local power controls media in China. *Media, Culture and Society*, *32*(6), 925–942. https://doi.org/10.1177/0163443710379665

Johansson, A. (2016). *Social media and politics in Indonesia*. Stockholm School of Economics Asia.

Kangei, L., Nyabul, P., & Muhenda, J. (2018). Habermasian deliberative democracy nuance: An enquiry. *International Journal of Advanced Scientific Research*, *3*(5), 45–53.

Lamb, K. (2019, September 5). Outcry as Indonesia seeks to arrest renowned West Papua rights lawyer. *The Guardian*. https://www.theguardian.com/world/2019/sep/05/outcry-as-indonesia-seeks-to-arrest-renowned-west-papua-rights-lawyer

Lim, M. (2017). Freedom to hate: Social media, algorithmic enclaves, and the rise of tribal nationalism in Indonesia. *Critical Asian Studies*, *49*(3), 411–427. https://doi.org/10.1080/14672715.2017.1341188

Lim, M. (2018). Disciplining dissent: Freedom, control, and digital activism in Southeast Asia. In R. Padawangi (Eds.), *Routledge handbook of urbanization in Southeast Asia* (pp. 478–494). Routledge

McQuail, D. (2010). *Mass communication theory*. SAGE.

Ministry of Communication and Information Technology of Indonesia. (2020a). *Undang-undang Republik Indonesia nomor 40 tahun 1999 tentang pers* [Law No. 40 of 1999 on the Press].

Ministry of Communication and Information Technology of Indonesia. (2020b). *Undang-undang Republik Indonesia nomor 32 tahun 2002 tentang penyiaran* [Broadcast Law No. 32/2002].

Ministry of Communication and Information Technology of Indonesia. (2020c). *Undang-undang Republik Indonesia nomor 19 tahun 2016 tentang informasi dan transaksi elektronik* [Information and Electronic Transaction Law No. 19/2016].

Muchtar, N., & Masduki. (2016). *The worlds of journalism study: Indonesia*. Worlds of Journalism Study.

Nadzir, I. (2020). *Hackers, doxers and influencers: The limits of political participation on social media.* Indonesia at Melbourne. https://indonesiaatmelbourne.unimelb.edu.au/hackers-doxers-and-influencers-the-limits-of-political-participation-on-social-media

Parikesit, B. (2020). The impact of surveillance on journalist activism. *Forum Ilmu Sosial*, *47*(2), 55–63. https://doi.org/10.15294/fis.v47i2.27057

Paterson, T. (2019). Indonesian cyberspace expansion: A double-edged sword. *Journal of Cyber Policy*, *4*(2), 216–234. https://doi.org/10.1080/23738871.2019.1627476

Pebrianto, F. (2020, August 21). Begini kronologi peretasan situs Tempo.co [The chronology of hacking to Tempo.co]. *Tempo*. https://nasional.tempo.co/read/1377884/begini-kronologi-peretasan-situs-tempo-co

Poell, T., & van Dijck, J. (2014). Social media and journalistic independence. In J. Bennett & N. Strange (Eds.), *Media independence: Working with freedom or working for free?* (pp. 182–201). Routledge.

Posetti, J. (2018). Combatting online abuse: When journalists and their sources are targeted. In J. Posetti & C. Ireton (Eds.), *Journalism, fake news and disinformation* (pp. 109–119). UNESCO.

Posetti, J., Aboulez, N., Bontcheva, K., Harrison, J., & Waisbord, S. (2020). *Online violence against women journalists*. UNESCO.

Power, T. (2018). Jokowi's authoritarian turn and Indonesia's democratic decline. *Bulletin of Indonesian Economic Studies*, *54*(3), 307–338. https://doi.org/10.1080/00074918.2018.1549918

Prabowo, H. (2019, April 15). Duduk perkara penghentian paksa nobar Sexy Killers di Indramayu [The chronology of forced stop to watch Sexy Killers movie in Indramayu]. *Tirto.id*. https://tirto.id/duduk-perkara-penghentian-paksa-nobar-sexy-killers-di-indramayu-dmaR

Putra, J., Manan, A., Madrin, S., & Murti, H. (2018). Doxing, persecution, and violence threatening journalists in Indonesia. In M. Hellema, C. Yi-Lan, & O. Motiwala (Eds.), *Freedom of expression under threat: Perspective from media and human right defenders in Asia* (pp. 17–22). FORUM ASIA.

Rahayu, K., Yogatama, B., & Patricia, S. (2019, September 24). Yang vokal yang diretas (1) [The hacked vocals]. *Kompas*. https://kompas.id/baca/utama/2019/09/24/yang-vokal-yang-diretas-1

Regan, P. (2012). *Regulating surveillance technologies*. Routledge.

Reporters Without Borders. (2018). *Online harassment of journalists*. https://rsf.org/en/news/rsf-publishes-report-online-harassment-journalists

Ruffini, P. (Ed.). (2018). *Underneath the autocrats: Southeast Asia media freedom report 2018—A report into impunity, journalist safety and working conditions*. International Federation of Journalists.

Rumata, V., & Sastrosubroto, A. S. (2018). Net-attack 2.0: Digital post-truth and its regulatory challenges in Indonesia. In R. Panuju (Ed.), *Proceedings of the International Conference of Communication Science Research (ICCSR 2018)* (pp. 116–120). Atlantis Press.

SAFEnet. (2021). Peluncuran laporan situasi hak-hak digital di Indonesia tahun 2020: Represi digital di tengah pandemi [Reports on human right situation in Indonesia 2020: Digital repression during pandemic]. https://id.safenet.or.id/2021/04/peluncuran-laporan-situasi-hak-hak-digital-di-indonesia-tahun-2020-represi-digital-di-tengah-pandemi/

Salam, F. (2020, August 24). Kronologi peretasan berita Tirto.id, dua artikel soal 'Obat Corona' [The chronology of hacking to Tirto.id news: Two articles about Corona medicine]. *Tirto.id*. https://tirto.id/kronologi-peretasan-berita-tirtoid-dua-artikel-soal-obat-corona-fZ2d

Saraswati, M. (2018). Social media and the political campaign industry in Indonesia. *Jurnal Komunikasi Ikatan Sarjana Komunikasi Indonesia*, *3*(1), 51–65.

Smith, D. (2017, January 23). Sean Spicer defends inauguration claim: 'Sometimes we can disagree with facts.' *The Guardian*. https://www.theguardian.com/us-news/2017/jan/23/sean-spicer-white-house-press-briefing-inauguration-alternative-facts

Sugiono, S. (2020). Fenomena industri buzzer di Indonesia: Sebuah kajian ekonomi politik media [The industry of buzzers in Indonesia: A political-economic analysis]. *Communicatus: Jurnal Ilmu Komunikasi*, *4*(1), 47–66. https://doi.org/10.15575/cjik.v4i1.7250

Tapsell, R. (2015). Indonesia's media oligarch and the 'Jokowi phenomenon.' *Indonesia*, *99*, 29–50. https://www.jstor.org/stable/10.5728/indonesia.99.0029

Tapsell, R. (2019, March 22). The polarization paradox in Indonesia's 2019 elections. *New Mandala*. https://www.newmandala.org/the-polarisation-paradox-in-indonesias-2019-elections

Tempo. (2020, December 10). The danger of political dynasties. *Tempo*. https://en.tempo.co/read/1413413/the-danger-of-political-dynasties

Wagstaff, J. (2014, March 28). Journalists, media under attack from hackers: Google researchers. *Reuters*. https://www.reuters.com/article/us-media-cybercrime-idUSBREA2R0EU20140328

Wahidin, K., & Ridwan, A. (2020, August 23). Kampanye Buzzer Ciptaker adalah blunder yang ironis [The buzzer campaign is an ironic blunder]. *Alinea*. https://www.alinea.id/politik/blunder-kampanye-buzzer-omnibus-law-ciptaker-b1ZSc9w8O

Walulya, G., & Nassanga, G. (2020). Democracy at stake: Self-censorship as a self-defence strategy for journal-

ists. *Media and Communication*, *8*(1), 5–14. https://doi.org/10.17645/mac.v8i1.2512

Warburton, E., & Aspinall, E. (2019). Explaining Indonesia's democratic regression: Structure, agency and popular opinion. *Contemporary Southeast Asia*, *41*(2), 255–285. https://doi.org/10.1355/cs41-2k

Yin, R. (2014). *Case study research: Design and methods* (5th ed.). SAGE.

About the Author

Masduki is an associate professor at the Department of Communication, Universitas Islam Indonesia, Yogyakarta, Indonesia. He earned his PhD at the Institute of Communication Studies and Media Research (IfKW), University of Munich, Germany (2019). He has published several books on the Indonesian broadcasting system and journalism. His articles have appeared in such scholarly journals as *GAZETTE*, *Journalism Studies*, *Journal of Digital Media and Policy* and *Media Asia*. Masduki has a particular interest in media policy, comparative media systems, broadcasting ethics, media activism, and journalism.

Media and Communication (ISSN: 2183–2439)
2021, Volume 9, Issue 4, Pages 62–72
https://doi.org/10.17645/mac.v9i4.4233

Article

Media Control and Citizen-Critical Publics in Russia: Are Some "Pigs" More Equal Than Others?

Rashid Gabdulhakov

Centre for Media and Journalism Studies, University of Groningen, The Netherlands; E-Mail: r.f.gabdulhakov@rug.nl

Submitted: 28 February 2021 | Accepted: 13 August 2021 | Published: 21 October 2021

Abstract
Amid the intensification of state control over the digital domain in Russia, what types of online activism are tolerated or even endorsed by the government and why? While entities such as the Anti-Corruption Foundation exposing the state are silenced through various tactics such as content blocking and removal, labelling the foundation a "foreign agent," and deeming it "extremist," other formations of citizens using digital media to expose "offences" performed by fellow citizens are operating freely. This article focuses on a vigilante group targeting "unscrupulous" merchants (often ethnic minorities and labour migrants) for the alleged sale of expired produce—the Hrushi Protiv. Supported by the government, Hrushi Protiv participants survey grocery chain stores and open-air markets for expired produce, a practice that often escalates into violence, while the process is filmed and edited to be uploaded to YouTube. These videos constitute unique media products that entertain the audience, ensuring the longevity of punitive measures via public exposure and shaming. Relying on Litvinenko and Toepfl's (2019) application of Toepfl's (2020) "leadership-critical," "policy-critical," and "uncritical" publics theory in the context of Russia, this article proposes a new category to describe state-approved digital vigilantes—citizen-critical publics. A collaboration with such publics allows the state to demonstrate a façade of civil society activism amid its silencing; while state-approved participants gain financial rewards and fame. Through Foucauldian discourse analysis, the article reveals that vulnerable groups such as labour migrants and ethnic minorities could fall victim to the side effects of this collaboration.

Keywords
authoritarian publics; digital vigilantism; Foucauldian discourse analysis; Hrushi Protiv; internet control; Russia; social justice

Issue
This article is part of the issue "Media Control Revisited: Challenges, Bottom-Up Resistance and Agency in the Digital Age" edited by Olga Dovbysh (University of Helsinki, Finland) and Esther Somfalvy (Research Centre for East European Studies at the University of Bremen, Germany).

1. Introduction

Since 2010, in grocery store chains and open-air food markets across Russia, one can witness people wearing full-body pig costumes surveilling shelves and counters for expired products. Such raids tend to escalate into verbal confrontations and physical violence between merchants and amateur inspectors who film everything and share edited videos on YouTube and other social media platforms, making them available to wide audiences. Beneath the pig outfits are former commissars of the pro-government youth movement Nashi (Ours) and other concerned citizens. Established in 2005, as a continuation of another pro-government organisation, Idushchiye Vmestye (Walking Together), Nashi, also known as Putin's Youth, was endorsed and sponsored by the state while actively supporting Vladimir Putin (see Hemment, 2012; Khalymonchik, 2016). Amid the decentralisation of Nashi and its consequent dissolution, several youth-led thematic activist formations emerged. One of the most prominent and still active projects among such groups is Hrushi Protiv. As per the group

itself, the title translates to "piggy against," although the literal translation is "piggies against." Transliteration from Cyrillic (Хрюши Против) into English can vary between *hrushi protiv*, *khrushi protiv*, *khryushi protiv*, and *khriushi protiv*; the name Hrushi Protiv will be used throughout this article based on the group's own use across its social media accounts.

In the case of Hrushi Protiv, retaliation turns into a form of entertainment, while participants acquire powers that turn them into grocery store reputation assassins. To conceptualise this form of citizen-led, digitally mediated justice provision, the article relies on the notion of digital vigilantism. Digital vigilantism can be defined as "direct online actions of targeted surveillance, dissuasion or punishment which tend to rely on public denunciation or an excess of unsolicited attention, and are carried out in the name of justice, order or safety" (Loveluck, 2019, p. 213). In digital vigilantism, visibility is the means and the ends of retaliation, as the very existence of publicity can have damaging effects when the names and locations of concerned businesses and the personal information of merchants are exposed to wide audiences. In this regard, being an entity with a unique online presence, Hrushi Protiv is not simply a case of conventional offline vigilantism being transferred to the online milieu; rather, it constitutes its own category of digitally mediated citizen-led justice—digital vigilantism.

Like Nashi, Hrushi Protiv is financially supported by the government and endorsed by Russia's top political leadership. On several occasions, Hrushi Protiv members have personally met with Vladimir Putin and former-president Dmitriy Medvedev, taking "selfies" and discussing social problems. Beyond verbal endorsements, the group has benefited from receiving as much as 21 million rubles (around 340,000 USD as per December 2019 conversion rates) in state grants (Public Verdict, n.d.). This intricate relationship of vigilantes and state leadership is especially intriguing given the wave of measures adopted by the government to regulate the digital domain (see, for instance, Lokot, 2020; Ognyanova, 2019; Vendil Pallin, 2016; Wijermars & Lehtisaari, 2021). Nevertheless, the liaison that the state has established with its loyal digitally savvy youth cannot be described one-dimensionally (Favarel-Garrigues & Shukan, 2019) and has been shown to be in flux. Some of the similar groups that emerged out of Nashi, such as the StopXam (Stop a Douchebag) movement, which counters bad parking, for instance, have had periods of both praise and condemnation by traditional media. Initially endorsed by the state in the same manner as Hrushi Protiv, StopXam may have crossed some boundaries of state trust when they started targeting important individuals. In 2016, for instance, the group publicly shamed and physically fought Russia's Olympic champion, consequently receiving a liquidation order issued by the Ministry of Justice (Gabdulhakov, 2019a).

This article addresses a complex relationship between digital vigilantes and the ruling elites amid the process of tightening state control applied on the digital domain, and in doing so, unveils various power hierarchies and webs of interests in state-citizen and citizen-to-citizen justice provision. Thus, the following main research questions are raised: Amid the ongoing crackdown on online self-expression in Russia, what types of citizen-initiated forms of online activism are tolerated and even endorsed by the government, and why? In addition, the article relies on three supporting questions: What are the motivations for participation in digital vigilantism? What are the impacts of digital vigilantism on those targeted? What role do platform affordances and regulation play in digital vigilantism?

The article first offers an overview of the scholarly discussion surrounding digital vigilantism and media systems in autocratic contexts generally and in Russia specifically. It proceeds with a description of its methodology and a presentation of the results following the application of Foucauldian discourse analysis on Hrushi Protiv's 20 most popular YouTube episodes. This is followed by a discussion of findings in reaction to the stated research questions. In conclusion, the article addresses theoretical implications and makes suggestions for future research.

2. Digital Vigilantism and Media Control in Russia

Connective actions, in which digital media serve as "organizing agents" for sharing "internalized or personalized ideas" (Bennett & Segerberg, 2012, pp. 752–753), have become a global phenomenon, at times capable of instigating social change through such movements as #MeToo (Mendes et al., 2018) or #BlackLivesMatter (Carney, 2016). Yet exposure and public shaming on social media can be characterised by different power dynamics, rendering respective layers of immunity and layers of vulnerability to targets, while participation can be both empowering and harmful (Gabdulhakov, 2019b, 2020).

Citizen-led justice manifested online can imply resistance against injustice and oppression as well as retaliation against the already vulnerable groups and individuals, such as minorities and migrants (Bjørgo & Mareš, 2019). Furthermore, vigilantes might use a façade cause to justify their actions while pursuing ulterior motives, be they political, ideological, financial, or other aspirations. Sometimes the motives are presented in bizarre combinations, such as the Serbian far-right nationalist vigilante group Levijatan (Leviathan), which claims to protect animal rights while engaging in "violent actions against Roma, LGBT and other 'enemies of Serbs' " (Colborne, 2020). Social justice and mob laws raise a number of questions related to legality, morality, effectiveness, and proportionality of citizen-to-citizen retaliation, especially when it comes to situations where, for whatever reason, authorised state services are replaced (or assisted) by vigilante forces.

2.1. Digital Vigilantism in Russia

After decades of scant scholarly attention to the notion of vigilantism, the phenomenon has recently gained momentum in the literature, with conceptual and empirical contributions featuring cases of divergent socio-political realities. Trottier (2017), for instance, offers a theoretical discussion on the role and impacts of visibility weaponisation in denunciatory acts. Moncada (2017), in turn, presents a classification of the varieties of vigilante practices and proposes core definitional dimensions for understanding the notion. With the focus on Russia's far-right, Kasra (2017) addresses the role of networked images in humiliation and socio-political control mechanisms in vigilante practices. Favarel-Garrigues (2019, 2021) elaborates on the entrepreneurial affordances of participants and their relationship with law enforcement. Loveluck (2019) develops a typology of digital vigilantism, relied upon in this article. Furthermore, the role of traditional media in facilitating digitally-mediated retaliation and rendering the phenomenon meaningful has been addressed in the ongoing debate (Gabdulhakov, 2019a). Despite the richness and depth of these contributions, the phenomenon requires further and continuous attention as approaches, environments, affordances, and nuances develop and evolve in real-time. Therefore, it is important to understand specific rules of engagement, respective power positions, benefits, and side effects of vigilante actions while also considering the unique affordances of social media and digital tools.

Loveluck (2019, p. 217) addresses the modes of coordination in digitally mediated vigilante practices and categorises them as ranging from "ad-hoc and loosely coordinated activities" to "pre-existing networks" that engage in "rehearsed collective efforts." In the quest for a typology of "online self-justice," he identifies four ideal types of digital vigilantism practices, namely: "flagging, investigating, hounding and organised leaking" (p. 214). Loveluck argues that in the process of flagging, targeting of the specific person involved, is avoided. Instead, the "low intensity" cases are meant to alert social media users by bringing to their attention instances of perceived norm-breaching (p. 217). Flagging via text and images is a global practice shared across social media platforms and political contexts. Unlike flagging, investigating implies naming the concerned target and a "collective effort" being made to investigate cases ranging from theft to more serious crimes and terrorist activities (Loveluck, 2019, p. 223). In this case, citizen-investigators are compared to the "web sleuths" who can provide their "technical expertise" in a given case (p. 224). Loveluck illustrates a complex dynamic between authorities, media, and web sleuths in which crowdsourced investigations do not terminate at the level of assisting police with the identification of criminals but can further evolve into digitally mediated harassment. "Hounding" takes matters to yet another level, referred to by Loveluck (2019, p. 227) as "the epitomy [sic] of digital vigilantism" it combines punitive intentions with investigations and mobilises participants against a specific target. Discreditation and public humiliation are the central aims in hounding. Finally, Loveluck presents "organised leaking" where participation is highly institutionalised and centred around the "documenting of problematic situations" and "the disclosure of confidential—and potentially incriminating—information" (p. 234). Examples of such organised groups include Russia's Anti-Corruption Foundation (FBK), whose activists investigate state corruption cases and publicise secret transactions of state officials. Some of the loudest investigations of the FBK shared on YouTube include the 2017 exposure of Russia's ex-president Dmitry Medvedev (Navalny, 2017) and the revelation of Russia's current president Vladimir Putin's riches (Navalny, 2021).

Activities of the FBK can serve as an example of what Rosenbaum and Sederberg (1974, p. 548) categorise as "regime control" vigilantism. In the absence of an official control mechanism that can be applied to the ruling elites, citizens take these duties into their own hands. Another group that can be classified as an example of organised leaking is Dissernet, a collective of academic enthusiasts who reveal plagiarism in doctoral dissertations. Operating in Russia and other former-Soviet states, Dissernet frequently targets state officials.

In the selected case study of Russia's Hrushi Protiv, hounding as a practice in digital vigilantism is most applicable. Much like other similar formations, activists of Hrushi Protiv indeed combine investigative approaches with practices of targeting specific businesses and individuals. Retaliation takes place not only in the form of verbal confrontations, physical fights, and destruction of produce; the targets and businesses that they represent can also suffer from long-lasting or even permanent reputational damage.

2.2. Media Control in Russia

Digitally mediated vigilante practices are part of the larger system combining political culture, social structures, media landscapes, and legal frameworks. Thus, it is necessary to elaborate on the milieu in which Hrushi Protiv operate. With the focus on Russia, this article seeks to address a context where the state's watchful gaze and control ambitions create a system that endorses some forms of online activism while cracking down on others. Having established a nearly totalitarian control over its traditional media sector, the government went after the digital domain with new legislation aimed at service providers, professional content creators, and individual users.

The waves of media landscape transformation in Russia are concurrent with major socio-political transformations in the country. Current processes demonstrate a past-oriented focus in terms of the Soviet-style information control strategies taking place in the new media landscape. These strategies include putting

pressure on service providers to filter content and share user data with the government and amending the legislation to criminalise certain forms of online self-expression, leading to large-scale arrests of social media users (Gabdulhakov, 2020; Lokot, 2020). This tendency for increased control is ongoing and reactionary since the government, for instance, also intervenes in the otherwise automated/algorithmic process of generating news feeds (Wijermars, 2021), among other approaches.

In their canonical work *Comparing Media Systems*, Hallin and Mancini (2004) propose three ideal types of media systems: democratic corporatist, liberal, and polarised pluralist. Each of the proposed systems is composed of dimensions, such as media market structure, political parallelism in news reporting, professionalisation of journalists, and the role of the state. Given the limited, West-centric case focus of Hallin and Mancini's original conceptualisation, Oates (2007) suggests that none of the three models can be applied to Russia, instead proposing the term "neo-Soviet" for the country's media model. Amid the multifaceted components that inform this model, such as bias, censorship, state, and commercial influences, mass media law, free speech protection, funding, media harassment, and violence against journalists, Oates offers a unique perspective by focusing on the position and the demands of the audience. Thus, when another major transformative wave in Russia's political, economic, and social sectors came about amid the collapse of the Soviet Union, the audience did not necessarily embrace the accompanying role of the media as a state critic. Akin to the Soviet media, which broadcasted based on national values, "giving the audience a sense of contentment and pride in their society," audiences in post-Soviet Russia, with a much wider variety of products at their disposal, valued mass media "as an institution that guides (rather than questions or undermines) the nation" (Oates, 2007, pp. 1295–1296). Public surveys conducted by Moscow-based Levada-Center (2018) demonstrate that in spite of the growth of the internet's popularity, the majority of people in Russia still rely on television as the main source of information and tend to trust it more than the internet.

Litvinenko and Toepfl (2019) react to another major political event in Russia's recent history, namely, the massive 2011–2013 protests for "Free and Fair Elections" (also known as protests on Bolotnaya Square). Dissent-curbing measures that followed these events once again reshaped Russia's media landscape. To understand the nature of this shift, Litvinenko and Toepfl (2019) rely on "authoritarian publics" theory (Toepfl, 2020) with the consideration of participants, environment and discursive practices and propose three types of publics—leadership-critical, policy-critical, and uncritical.

As such, several strategies have been adopted to counter the leadership-critical publics, following the mass protests in Moscow. Among these measures, Litvinenko and Toepfl (2019, pp. 232–233) identify "reining in discursive practices" via adopting of legal frameworks governing the digital domain and online self-expression, "shutting down environments" by blocking individual websites and platforms (blocking LinkedIn and attempting to block Telegram), and "intimidating participants" by limiting foreign media ownership, banning certain types of advertisement and replacing media owners with government-loyal elites. Relying on Schedler's (2013) "institutional gardening" concept to describe control measures, Litvinenko and Toepfl (2019, pp. 236) explain that policy-critical publics came out of the process of reshaping or "gardening" of leadership-critical publics. A vivid illustration of this reshaping is the metamorphosis of top leadership-critical news websites into policy-critical publics between 2012–2018 (Litvinenko & Toepfl, 2019, p. 235). Strategies shaping uncritical publics included recruiting civil servants, celebrities, active internet users, and paid PR workers known as "trolls" to exude vivid support for the political status quo "in novel Internet environments" (Litvinenko & Toepfl, 2019, pp. 235–236).

The intensity of the gardening of authoritarian publics in Russia is increasing. During the 2011 meeting with his supporters among online activists, then-president Medvedev called the Internet "an open space" and stated that even "things immoral in nature" have to be preserved online ("Dmitriy Medvedev vstretilsya so svoimi," 2011). The official rhetoric has shifted dramatically in one decade. During the 2021 meeting with Covid-19 pandemic-countering volunteer movement My Vmeste (We Are Together), President Putin called on the internet to "obey not even just laws, [as] formal legal rules, but also the moral laws of society," proceeding to label the internet as a source of "child pornography, child prostitution, promotion [and] distribution of drugs," a space where adolescents are "being pulled to the streets in order to misbehave there, [and] to fight with the police" ("Vstrecha s uchastnikami," 2021). Amid these shifts in perspectives, state critics are forced to strike a balance between reaching out to online audiences and managing personal risks that come along with such visibility (Lokot, 2018). At the same time, topics that can be subjected to public criticism are shrinking. By adopting strategic legislation and selectively applying the law, Russia's ruling elites continuously discourage citizens from criticizing the government and its policies (Lokot, 2020). Discussing, commenting, and even "liking" social media posts featuring taboo topics such as, for instance, anti-government protests, Crimea's annexation, or Russia's role in the Second World War can lead to legal scrutiny, fines, and prison sentences (Gabdulhakov, 2020). However, in this set of gardening mechanisms that shape authoritarian publics in Russia, it is still possible to engage in some forms of online activism, as is evident from the case of Hrushi Protiv. Building on Litvinenko and Toepfl's (2019) conceptualisation of leadership-critical, policy-critical and uncritical publics, this article proposes another category to describe the acts of state-approved digital

vigilantism—citizen-critical publics. Digital vigilantes can operate and target other citizens as long as these citizens do not represent or are not in any way connected to the ruling elites.

3. Methodology

Amid the wide variety of content analysis methods, the article relies on qualitative discourse analysis in Foucauldian terms. Describing the approach as one that "clearly refuses formalization" and has "no set rules," Arribas-Ayllon and Walkerdine propose selecting a corpus of statements, problematization, technologies, subject positions, and subjectification in order to conduct a Foucauldian discourse analysis (2008, p. 91). The authors identify five non-exhaustive types of corpora of statements suitable for a Foucauldian discourse analysis, namely: spatiality and social practice, political discourse, expert discourse, social interaction, and autobiographical accounts (2008, p. 100). Problematization may base itself on a response to the following questions: "Under what circumstances and by whom are aspects of human being rendered problematic, [and] according to what moral domains or judgement are these concerns allowed to circulate? What official discourses and counter-discourses render these problems visible and intelligible?" (Arribas-Ayllon & Walkerdine, 2008, p. 101). In Foucauldian discourse analysis, technologies are a concept that focuses on "power and self"—a type of " 'truth games' in which participants engage in conflict, competition and power" (p. 102). Subject positions in Foucauldian discourse analysis have to do with the moral order and the structure of rights and duties. Finally, subjectification refers to instances in which individuals self-regulate to "transform themselves in order to attain a certain state of happiness, purity, wisdom, perfection, or immortality" (Foucault, as cited in Arribas-Ayllon & Walkerdine, 2008, p. 103).

Commonly used in geography and psychology, Foucauldian discourse analysis is useful in addressing the aims of this interdisciplinary study, which incorporates elements of media studies and political science, by virtue of focusing on digital media affordances for citizen-led justice as well as the role of the state in media system formation and regulation. Applying a Foucauldian discourse analysis approach to the case of Hrushi Protiv in Russia, the article investigates how social hierarchies (Toelstede, 2020) inform current vigilante practices in the country and assesses the role of the official state position in rendering such practices meaningful amid the ongoing efforts to impose strict control over the digital domain.

Since 2010 Hrushi Protiv uploaded over 340 YouTube videos (as of 28 February 2021). As its corpus of statements, the article selected 20 of the most popular episodes in terms of the total number of views. When it comes to spatiality and social practice, Foucauldian discourse analysis allows for reliance on personal observations and ethnographic approaches. Online visibility is a weapon (Trottier, 2017) of punishment that Hrushi Protiv uses to harm its targets while simultaneously building its own brand and position as a justice provider in society. Given the significance of online artefacts in such practices, the article relied on netnographic approaches (Kozinets, 2015, 2019), which involved continuous online observation of Hrushi Protiv activities and content analysis of videos shared on the original Moscow-based group's YouTube channel.

Such observations were useful in understanding the nature and evolution of Hrushi Protiv raids. The author looked at the frequency of video uploads, the length of episodes, the number of views, comments, "likes" and "dislikes," and the titles of the episodes, which often resembled clickbait and yellow press headlines. In the initial phase, episodes were watched without a particular set of codes or categories in mind; the main goal was to get to know the group and to become familiar with its actions. As of 28 February 2021, Hrushi Protiv YouTube channel had 332,000 subscribers with 91,022,156 total video views, featuring 340 videos, the first of which was uploaded on 23 September 2010. Hrushi Protiv upload videos with varying frequency, but the practice is systematic, with at least one video released per month. The shortest video in the sample is 2 minutes and 31 seconds long, dedicated entirely to a fight between participants and targets at Moskvoretskaya produce base. The episode begins with a display of a link to a petition calling to ban migrants from retail work. The longest video is 26 minutes, featuring the raid of a grocery store staffed by ethnic minority employees. Out of 20 top videos, 11 were released in 2013, one in 2015, five in 2016, two in 2017, and one in 2019. This variation on the timeline of Hrushi Protiv activities suits Foucauldian discourse analysis' spatial focus.

The analysis additionally accounted for political and expert discourses, as Hrushi Protiv and similar formations that came out of Nashi have been endorsed by the state, while other manifestations of online citizen activism experience heavy state suppression. Political artefacts, in this respect, are public speeches, as well as formal and informal interactions between the government and participants. Expert discourses involve traditional media framing of participants and targets.

Social media affordances allow Hrushi Protiv to narrate their own autobiography, as it is communicated via online self-construction. The group and its members are relying on online communication modes in the process of defining the norms of morality and justice-provision methods while negotiating their own position in this equation. Inspired by methodological approaches of the grounded theory (Glaser & Strauss, 2017), this phase relied on an in-depth qualitative analysis (Altheide & Schneider, 2013) of Hrushi Protiv YouTube episodes with the focus on the positioning of self and respective framing of targets, police, and other actors appearing in the videos. YouTube itself constitutes a unique tool and a

stage for digital vigilantism, enabling both access to wide audiences, and money-making opportunities.

The author made several attempts to interview the founder as well as former and current members of Hrushi Protiv in Moscow and Saint Petersburg. In spite of an exchange of a few brief messages, participants did not agree to an interview. The author offered interview questions in written form, but the offer received no reaction. Why Hrushi Protiv members are reluctant to partake in an academic study is not particularly clear, but several reasons can be assumed. Perhaps, members had already been approached by one too many journalists and were either tired of giving interviews or saw no personal benefit in participating. The group is already rather well-known and can deliver any message they wish directly on their own social media pages and channels, without the involvement of third parties.

4. Hrushi Protiv on YouTube and Beyond

4.1. Corpus of Statements

Hrushi Protiv runs a website and has accounts on YouTube, VKontakte, Odnoklassniki, Facebook (the link to Facebook page provided on the official website and YouTube channel of the group was not functional in February 2021), Instagram, Twitter, Telegram, Live Journal, and TikTok. Social media profiles of Hrushi Protiv invite the viewers to support the project financially. Participants maintain an online store, where branded merchandise can be purchased. A separate website describing Hrushi Protiv as a "volunteer movement aimed at identifying trade in substandard products in stores" (Hrushi Protiv, n.d.) states that, in 2016, a branch was established in Belarus, making it an international group.

Most of the featured Hrushi Protiv episodes follow the same scenario in which activists equipped with video cameras enter stores and start loading the allegedly expired produce into shopping carts. Such acts lead to verbal and physical confrontations with store personnel which in some cases escalates into physical fighting. Content analysis revealed that violence (either featured in videos or promised in the titles) correlated with the popularity of these YouTube episodes. The most viewed episode was uploaded on 29.05.2019 and is called *Let's Step Outside*, a phrase commonly associated with an invitation to settle a conflict physically. Being 20 minutes and 15 seconds in length, it is one of the longer episodes of Hrushi Protiv with 3,665,938 views, 47,000 "likes," 11,000 "dislikes," and 21,439 comments (as of December 2019). In the episode, at least nine participants are shown entering the store. Grocery store personnel film participants with their phones while the latter raid the shelves. Verbal confrontations begin when personnel tell participants that filming is not allowed. Participants demand that targets explain the legal grounds for the prohibition of filming. The verbal back-and-forth continues for some time until the personnel give in and destroy the expired produce collected by the participants. Overall, 12 episodes out of 20 feature verbal and physical confrontation between participants and targets.

The signature trademark of Hrushi Protiv has been their full-body piggy outfit and is featured in half of the analysed episodes. Up until 2016, participants wore their piggy costumes consistently during the raids. Signature costumes made participants immediately visible and recognisable. In several videos, police ask participants where the costumes are, indicating popularity and recognition of the brand. For unclear reasons, starting from 2016, wearing piggy outfits became less consistent. Sometimes, activists are seen wearing branded shirts and hoodies featuring a piggy's head—the group's brand logo. Such merchandise is also available for sale in the group's online store. Other clothing items worn by participants include patriotic sports suits that read "Russia" across the back and hoodies with prints of Vladimir Putin in the military uniform of the commander in chief, emphasising the group's patriotic values and loyalty to the ruling regime.

In 10 out of the 20 episodes analysed, Hrushi Protiv target non-Slavic minorities. In another six episodes, the targets are mixed and include both non-Slavic minorities and the Slavs. Four episodes make no explicit reference to the ethnic backgrounds of targets. Thus, in 16 out of 20 episodes, a direct link between non-Slavic merchants and unscrupulousness in retail is emphasised. Hrushi Protiv openly expresses its prejudice towards labour migrants in Russia. In 2013–2014, participants called on their audience to sign a petition barring migrants from working in retail, an act suggestive of nationalist biases in these state-encouraged vigilante practices. One of the analysed episodes, titled *Hostages at Moskvoretskaya Produce Base*, features participants stating that "non-Russian employees run away when the police arrive" (Hrushi Protiv, 2017), emphasising both the "foreignness" of unscrupulous retailers as well as the potential illegality of "police-fearing" migrant workers.

Each episode uploaded by Hrushi Protiv is given a media-headline-like title, some of which are openly biased in terms of the ethnic background of the merchants, for instance: *Asian Showdown*, *We Don't Speak Russian*, *Tajiks Are Indignant*, *Migrants Beat Up Piggies*, etc. Other selected episodes contain such titles as *Real Jigits* (in some Turkic languages and in the Caucasus, the term *jigit* is used to describe brave young men), referring to the non-Slavic backgrounds of the targets, or *Moya Magazin* ("mine store"), which has an intentional grammatical mistake in the masculine noun, suggesting the target has a poor command of the Russian language. Overall, seven episode titles make explicit references to targeted retail workers' foreignness.

Police are featured in 11 of the 20 selected episodes. On three occasions, participants call the police to the site. In four cases, it is the targets who make such calls, and in five instances, it is not clear whose call the police

responded to. Police officers are generally passive, they register the names of all actors in both parties, collect the appeals and leave. In one episode, the activists are featured calling Russia's chief sanitary inspector, Gennady Onishchenko. In the video, Evgeniya Smorchkova apologises to Onishchenko for "calling again" and asks for help with a particular store that is not compliant with the demands. The next scene features the arrival of police officers at the store. The scenario in which participants directly call such a high-profile official (on more than one occasion) and ask for help, indicates the administrative capacities of the group, state endorsement, and support of their activities, and points to the power advantage that participants have over their targets.

Hrushi Protiv episodes occasionally feature informal leaders, such as celebrities. In one of the raids in the selected sample, participants are joined by a pop singer, member of a famous Russian boy band Ivanushki International. The artist does not engage in physical or verbal violence but is brought along to demonstrate the level of support and solidarity that Hrushi Protiv enjoy as citizen activists. Such informal endorsement once again stresses the unique capacities of participants and their ascendancy over targets.

4.2. Problematization

The internet and smart mobile devices have transformed the process of socialisation and surveillance at state-citizen and citizen-to-citizen levels in Russia. Numerous citizen formations establish thematic vigilante forces which target fellow citizens over alleged and perceived offences, such as bad parking, drinking, and smoking in public spaces, paedophilia (an accusation to which sexual minorities often fall target), drug dealing, and other "violations" of legal and moral boundaries. In some instances, no action is needed to attract the retaliation of vigilantes; simply being female (Avramov, 2019) or an ethnic minority (Chapman et al., 2018) is sufficient. In these realities of instrumentalisation of perceptions of morality for control of social order, Hrushi Protiv fulfil the function of an extension of the state, rather than being a collective of autonomous citizens. Much like the nostalgia for the Soviet-era media that communicated a sense of pride for the society, state-supported vigilante formations in Russia resemble various concerned groups of the past, such as the Tzarist and, subsequently, Soviet citizen-led justice provision formation Druzhina (Sokolov, 2019), the all-union Leninist young communist league Komsomol, or the system of comrades' courts that addressed minor mischiefs in breaching both legal and moral norms (Gabdulhakov, 2018).

4.3. Technologies

The case of Hrushi Protiv demonstrates how a citizen-led organisation can acquire legitimacy, recognition, and powers not only akin to those of official control entities (such as Russia's state sanitation service *Rospotrebnadzor*) but which also go beyond these entities in their technological savviness and retaliation approaches. Hrushi Protiv activities, in this regard, do not merely *flag* poor behaviour of their targets and cannot be compared to regular and widely practised consumer reviews, which inform fellow citizens about a particular business or product. Hrushi Protiv positions itself as a force operating between consumers and businesses as the former can report on the latter to participants. This position raises questions related to the possibility of intentional reputation damage upon orders from competitors of raided stores. What could stop "business A" from directly employing Hrushi Protiv or similar formations to expose a competing "business B"? One can only rely on the "good faith" of participants in this regard. At the same time, even with the assumed incorruptibility of participants, issues of legitimacy and proportionality of retaliation remain in question.

4.4. Subject Positions

Unlike a privately paid fine to state-controlled services due to misconduct, exposure on social media due to citizen-led retaliation brings about long-lasting reputational damage. Edited video materials uploaded by participants have the power to subject non-Russian targets to further scrutiny by police and immigration authorities. Given that the three Central Asian republics of Tajikistan, Kyrgyzstan, and Uzbekistan are highly remittance dependent (Bhutina, 2019), a labour migrant's job loss and/or deportation can lead to severe economic consequences for their families. The structure of power asymmetries (Toelstede, 2020) between participants and targets is informed by access to mass audiences on the one hand (and lack thereof), as well as social frustrations, ethnic, and national biases. Episodes tend to portray Hrushi Protiv and their targets as two fundamentally separate sociological clusters, with young participants being Slavic and grocery store or market personnel being comprised of non-Slavs.

4.5. Subjectification

Hrushi Protiv exemplifies a case where vigilant citizens acquire powers that give them wide social and media recognition. This visibly affords participants an almost TV persona stance. Hrushi Protiv even resembles the TV show, *Revizorro*, an adaptation of a Ukrainian show *Revizor*, airing on Russia's Pyatnitsa TV channel since 2014. The show's host exposes poor service provision practices in hotels and restaurants. The power of public exposure is so significant that businesses opt for collaboration with amateur controllers and sign agreements with Hrushi Protiv, promising to comply with imposed regulations (Hrushi Protiv, 2010) to avoid negative hype and reputation damage.

5. Discussion

Amid the intensification of state control over who can say what online in Russia, it is important to address the government's motives for supporting digital vigilantes. Hrushi Protiv and other similar formations are a product of the evolution of the Kremlin's youth policies and strategies that have undergone several overhauls. Nashi was formed as a national-patriotic movement to support the ruling elite and counter the opposition. Given that Nashi ceased to exist, former commissars of the movement needed a new project and issue-specific vigilante formations came into being. Having active and digitally savvy youth in its ranks is a convenient scenario for the regime, as long as this force does not turn against the patrons. The anti-migrant narratives of Hrushi Protiv, for instance, were handy in political campaigns constructed around the sentiments of threats coming from foreigners. However, in recent years, the Kremlin has adopted a harsher approach to relations with its former youth commissars. Active citizens are expected to turn into entities fully resembling Soviet-era loyal citizen squads extending the powers and omnipresence of the state.

Formations such as Hrushi Protiv are not threatening to the state unless they start targeting businesses that belong to the ruling elite. As long as certain boundaries are not crossed, the presence of such formations among the authoritarian publics allows for a display of an allegedly active civil society in realities where challenging state authority can carry large fines and lengthy prison sentences. Now in their 30s and having been engaged in the same vigilante practices for over a decade, some former Nashi activists have tried building political careers to various degrees of success. Perhaps the elites are allowing these citizen-critical publics to operate as a way of rewarding the once-loyal youth commissars for their support of the Putin–Medvedev tandem in the 2000s.

When it comes to motivation for participation in digital vigilantism, there are certain entrepreneurial interests (Favarel-Garrigues, 2019, 2021) as groups monetise YouTube channels, sell merchandise, advertise, ask for donations, and receive state grants to support their activities. In this sense, Hrushi Protiv is a formation with a hyper identity, simultaneously resembling citizen-led activism, a state-supported NGO, and a group of digitally-savvy entrepreneurs. Therefore, engagement in vigilante practices can afford participants financial and social benefits. Furthermore, endorsement by the state's highest authority affords legitimacy and provides a certain immunity when interacting with law enforcement.

What are the impacts of digital vigilantism on targets? Content analysis of the most viewed episodes shared by Hrushi Protiv on YouTube revealed ethnic and national biases in the group's activities. In most episodes, non-Russian or non-Slavic ethnic minorities are framed as untrustworthy, unscrupulous, aggressive, and violent. In fact, labour migrants are often in a fragile situation in terms of their legal status, difficult economic situation, and scarce employment opportunities in their home state. In their host state, then, they are even more vulnerable to online vigilantes amid a culture of xenophobia, police abuse, and a variety of other challenges. Sociological othering of non-Russian merchants might reflect on-the-ground offline frustrations, but such framing also creates discourses that shape and feed perceptions, leading to biased presumptions and stereotypes. In this regard, platforms such as YouTube become the central stage for such intra-citizen relations.

Beyond the questions of motives for participation and motives for state support of digital vigilantes, as well as the impact of such practices on individual and group targets, it is important to address platform affordances for digital vigilantism. Platforms such as YouTube allow participants to acquire a large following and generate an income via monetisation and advertising. Participants are able to create discourse through their own channels by editing the videos and accompanying comments. As such, YouTube enables an environment in which digital vigilantism is manifest. Such manifestation, however, is taking place on uneven grounds and at the crossroads of various interests. For instance, citizen-critical content featuring inter-ethnic hostility, such as Hrushi Protiv's calls for a ban on migrants working in retail, can freely circulate the internet, while state-critical and policy-critical content is deemed extremist.

Several important aspects come to the surface here. The first has to do with political regimes and internet governance. When pressure is put on platforms to moderate content, there is a threat that select voices challenging the political status quo will be muted, as is evident in the case of Russia. When the opposition-led FBK exposed Russia's deputy prime minister for accepting a bribe from a prominent oligarch, the government put pressure on platforms, and Facebook's daughter company Instagram complied with the requests to remove posts over privacy concerns (Nechepurenko, 2018). The fine line between the right to privacy and power abuse for covering up corruption is blurred in this case. This example demonstrates the spill-over effect of biased institutions on social media platforms and the selective application of the law. In this governance environment, both domestic and global social networking corporations can fall target to invasive state control aimed at serving the interests of the ruling elites.

Those with political and financial power seem to continue enjoying the privileges and immunities online, while the powerless, such as migrants, ethnic, sexual and other minorities, political activists, women, journalists, are vulnerable. The role of platforms in the facilitation of select hounding (Loveluck, 2019) practices and the power and logic of removal of undesired content need to be addressed at both analytical and policy levels. At the same time, an important question to ask is: Would critical publics in Russia benefit from any state regulation of platforms in a context where ruling elites are able to actively silence critical voices?

6. Conclusion

This article provided a detailed account of Hrushi Protiv activists operating across and beyond Russia. Having addressed the peculiarities of vigilante practices in Russia, the article demonstrated that the state plays a central role in (dis)approving digitally mediated citizen-led initiatives as part of its strategies for the gardening of authoritarian publics (Litvinenko & Toepfl, 2019). Through the selected case, this article offers a detailed account of how vigilante formations such as Hrushi Protiv weaponise hounding (Loveluck, 2019) to acquire financial (sometimes political) and other benefits from their activities. By being loyal to the ruling elites and not crossing boundaries that could potentially harm them, formations such as Hrushi Protiv are allowed to operate in what are otherwise tightly controlled digital and public domains. The government benefits from such citizen-critical publics. First of all, the blame is taken off the political elites and policies. Citizen-critical publics elevate on-the-ground unscrupulousness, as opposed to challenging the system itself. At the same time, amid control intensification, the government may aim to appear less repressive than it actually is by demonstrating a façade of an active civil society in the country.

In Russia and elsewhere, digital vigilantism is practised and perceived as a form of entertainment akin to reality TV shows, with each episode carefully edited and professionally arranged with catchy titles, music, and other strategies, such as the featuring of celebrity guests. It is evident that vigilante activities constitute a reflection of on-the-ground societal frustrations and tensions. Traffic jams and poor parking, cheated customers, xenophobia, homophobia, labour migration, and other "hot" societal issues in Russia are picked up and instrumentalised by vigilantes who step in and turn battling against perceived injustices into a spectacle. In this case, the citizen-critical focus of YouTube videos is not only safe but is arguably beneficial for the ruling elites amid their strategy to discourage leadership-critical and policy-critical discourse.

Further research on the subject could focus on comments left under YouTube episodes to measure audience perceptions of citizen-critical publics, although it should be noted that channel owners can mute and otherwise moderate reactions. Comparative studies focusing on formations similar to Hrushi Protiv in other socio-political and media contexts would help advance theoretical boundaries of the phenomenon of digital vigilantism and media system models.

Acknowledgments

This work was supported by the Netherlands Organisation for Scientific Research (NWO; grant no. 276-45-004).

Conflict of Interests

The author declares no conflict of interest.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Altheide, D. L., & Schneider, C. J. (2013). *Qualitative media analysis*. SAGE.

Arribas-Ayllon, M., & Walkerdine, V. (2008). Foucauldian discourse analysis. In C. Willig & W. Stainton-Rogers (Eds.), *The SAGE handbook of qualitative research in psychology* (pp. 91–108). SAGE. https://www-doi-org.eur.idm.oclc.org/10.4135/9781848607927

Avramov, K. (2019, March 27). Russia's virtual moral police: Toxic subculture in pursuit of purity. *The Globe Post*. https://theglobepost.com/2019/03/27/russia-male-state

Bennett, W. L., & Segerberg, A. (2012). The logic of connective action: Digital media and the personalization of contentious politics. *Information Communication and Society*, *15*(5), 739–768. https://doi.org/10.1080/1369118X.2012.670661

Bhutina, S. (2019, May 23). Russian remittances to Central Asia rise again. When Russia hurts, Central Asians feel the pain. Are remittances a boon or a bane? *Eurasianet*. https://eurasianet.org/russian-remittances-to-central-asia-rise-again

Bjørgo, T., & Mareš, M. (Eds.). (2019). *Vigilantism against migrants and minorities*. Routledge.

Carney, N. (2016). All lives matter, but so does race: Black lives matter and the evolving role of social media. *Humanity & Society*, *40*(2), 180–199. https://doi-org.eur.idm.oclc.org/10.1177/0160597616643868

Chapman, H. S., Marquardt, K. L., Herrera, Y. M., & Gerber, T. P. (2018). Xenophobia on the rise? Temporal and regional trends in xenophobic attitudes in Russia. *Comparative Politics*, *50*(3), 381–394. https://www.jstor.org/stable/26532692

Colborne, M. (2020, June 18). Levijatan: Serbian animal rights vigilantes go to the Polls. *Bellingcat*. https://www.bellingcat.com/news/2020/06/18/levijatan-serbian-animal-rights-vigilantes-go-to-the-polls/?amp=1

Dmitriy Medvedev vstretilsya so svoimi storonnikami—Predstavitelyami setevykh soobshchestv [Dmitry Medvedev met with his supporters—Representatives of online communities]. (2011, November 9). *The Kremlin*. http://kremlin.ru/events/president/news/13443

Favarel-Garrigues, G. (2019). Digital vigilantism and anti-paedophile activism in Russia: Between civic involvement in law enforcement, moral policing and business venture. *Global Crime*, *21*(3-4), 306–326.

https://doi.org/10.1080/17440572.2019.1676738
Favarel-Garrigues, G. (2021). 'Vigilante shows' and law enforcement in Russia. *Europe-Asia Studies*, *73*(1), 221–242. https://doi.org/10.1080/09668136.2020.1862061
Favarel-Garrigues, G., & Shukan, I. (2019). Perspectives on post-Soviet vigilantism. Introduction. *Laboratorium: Russian Review of Social Research*, *11*(3), 4–15.
Gabdulhakov, R. (2018). Citizen-led justice in post-communist Russia: From comrades' courts to dot-comrade vigilantism. *Surveillance & Society*, *16*(3), 314–331. https://doi.org/10.24908/ss.v16i3.6952
Gabdulhakov, R. (2019a). Heroes or hooligans? Media portrayal of StopXam (Stop a Douchebag) vigilantes in Russia. *Laboratorium: Russian Review of Social Research*, *11*(3), 16–45. https://doi.org/10.25285/2078-1938-2019-11-3-16-45
Gabdulhakov, R. (2019b). In the bullseye of vigilantes: Mediated vulnerabilities of Kyrgyz labour migrants in Russia. *Media and Communication*, *7*(2), 230–241. https://doi.org/10.17645/mac.v7i2.1927
Gabdulhakov, R. (2020). (Con)trolling the web: Social media user arrests, state-supported vigilantism and citizen counter-forces in Russia. *Global Crime*, *21*(3/4), 283–305. https://doi.org/10.1080/17440572.2020.1719836
Glaser, B. G., & Strauss, A. L. (2017). *Discovery of grounded theory: Strategies for qualitative research*. Taylor & Francis Group.
Hallin, D. C., & Mancini, P. (2004). *Comparing media systems: Three models of media and Politics*. Cambridge University Press.
Hemment, J. (2012). Nashi, youth voluntarism, and Potemkin NGOs: Making sense of civil society in post-soviet Russia. *Slavic Review*, *71*(2), 234–260. https://doi.org/10.1017/s0037677900013607
Hrushi Protiv. (n.d.). *About*. https://хрюши.рф/about
Hrushi Protiv. (2010). *Sovmestnaya press-konferentsiya kompanii KH5 i Khryush v "RIA Novosti" 6 dekabrya* [A joint press conference of X5 Company and Hrushi at "RIA Novosti" on 6 December]. https://hrushi-protiv.livejournal.com/22960.html
Hrushi Protiv. (2017, July 19). *Hrushi protiv—Zalozhniki na moskvoretskoy* baze [Piggy against—Hostages at the Moskvoretskaya base] [Video]. https://www.youtube.com/watch?v=LJ4P8kiwyho
Kasra, M. (2017). Vigilantism, public shaming, and social media hegemony: The role of digital-networked images in humiliation and sociopolitical control. *The Communication Review*, *20*(3), 172–188. https://doi-org.eur.idm.oclc.org/10.1080/10714421.2017.1343068
Khalymonchik, N. (2016). The quest for ideal youth in Putin's Russia II. The search for distinctive conformism in the political communication of Nashi, 2005–2009. *Europe–Asia Studies*, *68*(6), 1089–1091. https://doi.org/10.1080/09668136.2016.1202534
Kozinets, R. (2019). *Netnography: The essential guide to qualitative social media research* (3rd ed.). SAGE.
Kozinets, R. V. (2015). *Netnography: Redefined* (2nd ed.). SAGE.
Levada-Center. (2018, October 12). *Channels of information*. https://www.levada.ru/en/2018/10/12/channels-of-information
Litvinenko, A., & Toepfl, F. (2019). The "gardening" of an authoritarian public at large: How Russia's ruling elites transformed the country's media landscape after the 2011/12 protests "for fair elections." *Publizistik*, *64*(2), 225–240. https://doi.org/10.1007/s11616-019-00486-2
Lokot, T. (2018). Be safe or be seen? How Russian activists negotiate visibility and security in online resistance practices. *Surveillance & Society*, *16*(3), 332–346. https://doi.org/10.24908/ss.v16i3.6967
Lokot, T. (2020). Articulating networked citizenship on the Russian internet: A case for competing affordances. *Social Media Society*, *6*(4), https://doi.org/10.1177/2056305120984459
Loveluck, B. (2019). The many shades of digital vigilantism. A typology of online self-justice. *Global Crime*, *21*(3/4), 213–241. https://doi.org/10.1080/17440572.2019.1614444
Mendes, K., Ringrose, J., & Keller, J. (2018). #metoo and the promise and pitfalls of challenging rape culture through digital feminist activism. *European Journal of Women's Studies*, *25*(2), 236–246.
Moncada, E. (2017). Varieties of vigilantism: Conceptual discord, meaning and strategies. *Global Crime*, *18*(4), 403–423. https://doi.org/10.1080/17440572.2017.1374183
Navalny, A. (2017, March 2). *On vam ne Dimon* [Don't call him Dimon; Video]. https://www.youtube.com/watch?v=qrwlk7_GF9g
Navalny, A. (2021, January 19). *Dvorets dlya Putina. Istoriya samoy bol'shoy vzyatki* [Palace for Putin. The history of the biggest bribe; Video]. https://www.youtube.com/watch?v=ipAnwilMncI&t=197s
Nechepurenko, I. (2018, February 15). Russia locks Aleksei Navalny's website, after his inquiry into an oligarch. *The New York Times*. https://www.nytimes.com/2018/02/15/world/europe/russia-navalny-instagram-youtube-deripaska.html
Oates, S. (2007). The neo-Soviet model of the media. *Europe–Asia Studies*, *59*(8), 1279–1279. https://doi.org/10.1080/09668130701655150
Ognyanova, K. (2019). In Putin's Russia, information has you: Media control and internet censorship in the Russian Federation. In Information Resources Management Association (Ed.), *Censorship, surveillance, and privacy: Concepts, methodologies, tools, and applications* (pp. 1768–1786). IGI Gobal.
Public Verdict. (n.d.). *Dos'ye Hrushi Protiv* [Dossier: Piggy Against]. http://vigilant.myverdict.org/files/pigs
Rosenbaum, H. J., & Sederberg, P. C. (1974). Vigilantism: An analysis of establishment violence. *Comparative Politics*, *6*(4), 541–570. https://doi.org/

10.2307/421337
Schedler, A. (2013). *The politics of uncertainty: Sustaining and subverting electoral authoritarianism*. Oxford University Press.
Sokolov, N. (2019, June 25). Agressivnykh blogerov i aktivistov predlozheno sdelat' druzhinnikami [It is proposed to turn the aggressive bloggers and activists proposed into druzhina]. *Vesti*. https://www.vesti.ru/videos/show/vid/802189
Toelstede, B. (2020). Social hierarchies in democracies and authoritarianism: The balance between power asymmetries and principal-agent chains. *Rationality and Society*, *32*(3), 334–366. https://doi.org/10.1177/1043463120904051
Toepfl, F. (2020). Comparing authoritarian publics: The benefits and risks of three types of publics for autocrats. *Communication Theory*, *30*(2), 105–125. https://doi.org/10.1093/ct/qtz015
Trottier, D. (2017). Digital vigilantism as weaponisation of visibility. *Philosophy and Technology*, *30*(1), 55–72. https://doi.org/10.1007/s13347-016-0216-4
Vendil Pallin, C. (2016). Internet control through ownership: The case of Russia. *Post-Soviet Affairs*, *33*(1), 16–33. https://doi.org/10.1080/1060586X.2015.1121712
Vstrecha s uchastnikami aktsii 'My Vmeste' [A meeting with participants of 'We Are Together']. (2021, March 4). *The Kremlin*. http://kremlin.ru/events/president/news/65096
Wijermars, M. (2021). Russia's law 'on news aggregators': Control the news feed, control the news? *Journalism*. Advance online publication. https://doi.org/10.1177/1464884921990917
Wijermars, M., & Lehtisaari, K. (2021). *Freedom of expression in Russia's new mediasphere* (1st ed.). Routledge.

About the Author

Rashid Gabdulhakov is an assistant professor at the Centre for Media and Journalism Studies at the University of Groningen, the Netherlands. Rashid received his PhD (cum laude) in Media and Communication from Erasmus University Rotterdam, the Netherlands. He holds a MA of Advanced Studies degree in International and European Security from the University of Geneva and the Geneva Centre for Security Policy, Switzerland; and a Master of Arts degree in Politics and Security from the Organisation for Security and Cooperation in Europe (OSCE) Academy in Bishkek, Kyrgyz Republic. For more information and a full CV, please visit: www.plovism.com

Media and Communication (ISSN: 2183–2439)
2021, Volume 9, Issue 4, Pages 73–81
https://doi.org/10.17645/mac.v9i4.4238

Article

Boundary Control as Gatekeeping in Facebook Groups

Sanna Malinen

Department of Social Research, University of Turku, Finland; E-Mail: sanna.malinen@utu.fi

Submitted: 28 February 2021 | Accepted: 21 May 2021 | Published: 21 October 2021

Abstract
Facebook groups host user-created communities on Facebook's global platform, and their administrative structure consists of members, volunteer moderators, and governance mechanisms developed by the platform itself. This study presents the viewpoints of volunteers who moderate groups on Facebook that are dedicated to political discussion. It sheds light on how they enact their day-to-day moderation work, from platform administration to group membership, while acknowledging the demands that come from both these tasks. As volunteer moderators make key decisions about content, their work significantly shapes public discussion in their groups. Using data obtained from 15 face-to-face interviews, this qualitative study sheds light on volunteer moderation as a means of media control in complex digital networks. The findings show that moderation concerns not just the removal of content or contacts but, most importantly, it is about protecting group norms by controlling who has the access to the group. Facebook's volunteer moderators have power not only to guide discussion but, above all, to decide who can participate in it, which makes them important gatekeepers of the digital public sphere.

Keywords
Facebook; Facebook groups; gatekeeping; moderation; platforms; political discussions; social networks

Issue
This article is part of the issue "Media Control Revisited: Challenges, Bottom-Up Resistance, and Agency in the Digital Age" edited by Olga Dovbysh (University of Helsinki, Finland) and Esther Somfalvy (Research Centre for East European Studies at the University of Bremen, Germany).

1. Introduction

Social media platforms provide users with vibrant spaces for public discussions across a wide range of topics. By giving individuals the power to network independent of institutions, social media increases collective action and accountability in society (Bennett & Segerberg, 2012; Dutton, 2009; Gustafsson, 2012; Kushin & Kitchener, 2009). In particular, Facebook groups have been identified as a significant arena for citizen engagement as they allow discussion of common interests and goals (Park et al., 2009), and group identity and self-efficacy to be built in relation to participation (Gustafsson, 2012). Currently, Facebook hosts a large number of politically motivated user groups created by political actors and civil activists that reach wide audiences (Gustafsson, 2012; Park et al., 2009; Warren et al., 2014). Previous studies have pointed out that groups on Facebook can promote societal change and provide users with a channel for expressing counter-discourses to the dominant public voice (Gachau, 2016; Pruchniewska, 2019; Sormanen & Dutton, 2015). However, there is a darker side to social media, and research has pointed out the harmful effects of such platforms on political, economic, and social life, particularly due to the widespread dissemination of misinformation and hate speech through them (Bessi et al., 2016; Del Vicario et al., 2016; Gagliardone et al., 2015).

Social media platforms have the power to promote, delete, and hide content produced by users, making them an important means to shape public discussion (Gillespie, 2018; Gorwa, 2019; Myers West, 2017). Even though they were originally created for facilitating social activity between people and increasing the circulation of user-generated content, they need to be moderated to keep discussions civil and law-abiding. Previous studies have suggested that users are unaware of the moderation policies of platforms and the underlying logic

of these, and these policies are intentionally being kept guarded in order to maintain a sense of openness and freedom that any form of moderation and content control would typically be strongly against (Gillespie, 2018; Roberts, 2016). As Roberts (2016) argued, social media companies want to give their users the impression that content appears on the site simply "in some kind of natural, organic way" (p. 9), and therefore they intentionally obscure human decision-making processes behind moderation. In other words, commercial content moderation is successful when it is invisible as it is not intended to leave any traces (Gillespie, 2018; Roberts, 2016).

Platforms use various moderation strategies for promoting and discarding content. Recently, algorithmic moderation and so-called "filter bubbles" (Pariser, 2011) have received plenty of scholarly attention; but still, less is known about how user-driven modes of content control are organized. Kalsnes and Ihlebæk (2021) argued that user-driven moderation should be viewed as political because volunteer moderators choose to give visibility to some views while hiding others. Particularly when moderation decisions lack transparency, they have serious consequences for participation in public debate (Kalsnes & Ihlebæk, 2021). So far, the obscure nature of social media platforms has made it difficult to study their moderation systems thoroughly (Jhaver et al., 2019; Langlois et al., 2009).

Facebook groups are local, user-created groups hosted by a global platform, which makes their moderation structure complex. Groups' founders can create and enact their own governance policies, and therefore the rules and moderation practices of individual groups vary greatly throughout the platform. This study focuses on one—and perhaps the most visible—aspect of how these groups are moderated: volunteer moderators who make decisions about acceptable content on a daily basis. It investigates how these moderators create and enact moderating philosophies as intermediaries between group members and the platform. Since not much is known about the work of volunteer moderators, it is important to shed more light on how they shape the visibility of political views in networked social spaces. This study relies on network gatekeeping theory introduced by Barzilai-Nahon (2008) and looks into social dynamics between the stakeholders in the political communities on Facebook groups.

2. Theoretical Background

2.1. Content Moderation

Ever since the emergence of online communities, the way these are moderated has mostly been the responsibility of their founders and key members (Kalsnes & Ihlebæk, 2021). Over the years, scholars have debated whether volunteer moderation is emotionally demanding and labor-intensive unpaid work conducted for the benefit of the companies that run the platforms (e.g., Terranova, 2000) or an organic part of community management and development (Seering et al., 2019). Moderators have a key role in determining which content is published and what is removed, and these decisions shape our public discourse (Gillespie, 2018; Jhaver et al., 2019). According to Kalsnes and Ihlebæk (2021), the role of moderators has recently grown to become even more important with the growing prevalence of uncivil online behavior, such as harassment and hate speech that poses a threat to democracy.

Major social media companies have developed moderation strategies for monitoring user-generated content on platforms. However, volunteer users are still the most effective moderators because they understand group norms, are strongly committed to their communities, and derive personal meaning from their moderation work (Gillespie, 2018; Seering et al., 2019). Prior research has shown that volunteer moderators tend to engage personally in the moderation process and view it as a means of growth for both themselves and their communities (Seering et al., 2019). When moderation decisions are left to algorithms or paid moderation teams, communities and their human moderators miss opportunities for guiding discussion and reflecting the values behind it (Ruckenstein & Turunen, 2020; Seering et al., 2019).

User communities hosted by social media platforms differ from traditional, self-governing online communities in terms of their structure. On platforms such as Facebook or Reddit, users can create their own subgroups and develop specific local policies alongside the platforms' site-wide rules and terms of use. This complex approach to policy is well exemplified in Facebook groups, whereby users navigate between Facebook's own community norms, multiple individually tailored, community-devised rules, and implicit cultural codes of conduct. Operating in a multi-layer system with rules derived from a range of sources can be confusing for users (Fiesler et al., 2015). The rules of local groups not only vary across groups, but they can also be rather vague. According to Dovbysh (2021), rules of moderation are individually constructed by group owners who imitate journalistic practices, although they lack professional norms, as was found in the study on Russian Vkontakte groups.

Facebook groups combine both commercial governing mechanisms developed by the platform and self-governance by group members. These two modes of moderation are clearly different in terms of their impact on a group's social dynamics. Volunteer moderators work from the bottom up, whereas commercial content moderation is directed from the top down, obeying policy and norms set by company. A study by Seering et al. (2019) showed that while company-driven moderation strategies view anti-normative behavior as something that should be removed or banned, volunteer moderators tend to personally engage with community members and view such interaction as an opportunity for growth for the whole community. Some scholars have emphasized this continuous interaction as a

centerpiece for community development and conceptualized volunteer moderation as an ongoing negotiation in which the meaning of moderation is continuously defined and explained amongst stakeholders such as the platform, community, and fellow moderators (Gillespie, 2018; Matias, 2019; Seering et al., 2019). This implies that community guidelines are not fixed and can evolve over time as a result of a company's self-perception and the demands of users; in other words, they are consequent to the negotiation process (Myers West, 2017).

Prior studies have identified fairness as a key element of successful moderation as users' reaction to moderation is likely to depend on whether they feel it is done fairly (Jhaver et al., 2019; Myers West, 2017). If there is confusion about the reasons for moderation or feelings of being treated unfairly, users who have experienced moderation can become frustrated. One way for them to deal with this frustration and confusion is by developing their own theories for content takedown (Jhaver et al., 2019; Myers West, 2017). In particular, hidden commercial content moderation creates tensions between users and the platform: Frustrated users may turn against platforms through collective protests with the aim of raising the visibility of content that the platform has hidden from them (Gillespie, 2018; Myers West, 2017). In the commercial moderation system, users remain absent, and they are only given the role of laborers who can report content they deem objectionable (Myers West, 2017).

2.2. Gatekeeping in Social Networks

For decades, scholars of media and communication have applied a theory of gatekeeping to describe content selection in the media environment and ascribed the term "gatekeeper" to persons who have a role in carrying out this selection. Barzilai-Nahon (2008) has addressed the need for updating traditional gatekeeping theories to fit better in the context of digital networks. According to her, traditional theories view gatekeeping as a selection process based on a gatekeeper's individual characteristics and position of power, while dynamics and relationships between stakeholders are left unconsidered. This reduces gatekeeping to simply a one-way direction and top-down process, which is an inadequate way to describe it in the context of information networks with multiple gates and channels for spreading information (Barzilai-Nahon, 2008). In the context of this study of groups on Facebook, this complexity of information flow is seen in volunteer moderators' ability to control only information within their own communities. Network gatekeeping theory presents three main goals: "locking-in" of gated users inside the gatekeeper's network; protecting established communities from unwanted entry from outside; and maintaining ongoing activities within network boundaries without disturbances (Barzilai-Nahon, 2008). All three goals point to outsiders being the main threat for communities.

Contrary to the traditional literature, which has conceived the gatekeeper as having complete power in information production and dissemination, Barzilai-Nahon (2008) saw a dynamic relationship between the gatekeeper and the gated that forms through frequent, enduring, and direct exchange. The gated are not viewed as passive nor powerless in this process; they too can have power and exercise it. Contrary to traditional media settings, non-elite members can become prominent in gatekeeping in the networked context as well, influencing what is being discussed and how it is done. This is particularly evident in cases of mass movements and uprisings, when ordinary users play a significant role in raising topics to prominence and elevating others to higher status through active gatekeeping (Meraz & Papacharissi, 2013).

Collaborative and networked modes of action were expected to lead to flatter and less hierarchical organizational forms (Bennett & Segerberg, 2012). However, there is evidence that power structures of user-driven communities can be rather oligarchic, so that some individuals gain a more privileged position and exert their authority onto others (Keegan & Gergle, 2010; Shaw & Hill, 2014). As shown by a study of Wikipedia, elite users are in a position to select and remove content, but they also have to accept contributions from non-elite users in order to keep content flowing (Keegan & Gergle, 2010). Similarly, there are elite users with privileges to restrict others in software wiki communities, although they can hinder community development when they use this authority to promote their own agendas over the interests of the community as a whole (Shaw & Hill, 2014). Scholars have presented a range of views on participatory structures of online communities and how these structures are associated with moderation (Keegan & Gergle, 2010; Matias, 2019; Seering et al., 2019; Shaw & Hill, 2014). The key questions here are how should privileged members exert their power over ordinary members, and should they restrict some users to maintain the harmony of the community? In order to survive, online communities need to self-regulate, members must conform to norms by monitoring their own behavior, and those who violate these norms should be punished (Honeycutt, 2005).

The network gatekeeping theory recognizes that the stakeholders involved in gatekeeping are not equally powerful, and some attributes, such as political power, information production ability, or relationship with the gatekeeper, can lead to greater salience in the network (Barzilai-Nahon, 2008). As Myers West (2017) argued, visibility is the most effective way to gain political power on social media in a networked environment, and users without the power to influence platforms' moderation policy can fight unfair moderating decisions by giving prominence to what platform has hidden.

The starting point of this study is that social media users are not just passive receivers of information; instead, they can actively construct their political

environment on social media by building networks and tailoring their information flows. Relying on network gatekeeping theory and its two main components, network gatekeeping identification and network gatekeeping salience, this study aims to investigate gatekeeping practices and goals in political Facebook groups (RQ1) and analyze power dynamics between the gatekeepers and the gated (RQ2). The term "salience" refers to the degree to which gatekeepers give priority to the gated. Therefore, this study examines if there are any differences in members' positions of power (RQ3), so that some group members are more influential and therefore gain more visibility for their views than others manage to do.

3. Method and Data

This qualitative study uses data obtained from 15 semi-structured interviews with Facebook group moderators as research data. The face-to-face interviews were conducted between December 2019 and February 2020 in Finland. The informants were selected first by searching for active Finnish Facebook discussion groups labeled as political or societal. Then persons named as moderators or administrators of these groups were identified through each group's public page and contacted personally via Messenger. Initially, interview requests were sent to 20 individuals, of whom five either declined or did not see the invitation. The interviews were recorded and transcribed, and the duration of the voice files varied from 58 to 170 minutes.

This study uses semi-structured interviews because of the flexibility of this format. In addition to predetermined research themes, it allows other relevant themes to develop throughout the interviews (Choak, 2012). Therefore, it can bring out new and unexpected results and allow the study to take new directions. Data analysis followed thematic analysis, which is a process of identifying patterns and themes within the data (Creswell, 2013). The analysis first focused on gaining a detailed understanding of moderation practices and the interviewees' experiences of their roles. Following the procedure described by Creswell (2013), the text was first classified into codes and then into broader themes. Each of the themes were interpreted in terms of their meaning in respect to the research questions.

This study focuses on groups dedicated to political discussion for three reasons. First, previous research showed that tensions between users of social media networks tend to arise particularly when discussion is connected to politics (Zhu et al., 2017). Second, political beliefs and attitudes are found to drive selectivity in subsequent information processing (Taber & Lodge, 2006). Third, in a political context, information control reflects the state of power relations between stakeholders who aim to achieve their political goals (Barzilai-Nahon, 2008). Prior work has thus suggested that content selection and moderation are very likely to occur in politically motivated social media discussions in comparison with other topics.

4. Findings

4.1. Moderation as Boundary Control

"You don't want someone stupid at a good party" (Interviewee 6). As shown in the quote, the interviewees indicated that screening applications for group membership is an essential aspect of moderation. Many moderators reported having developed specific checklists to evaluate who would become a suitable member and contributor and who is applying to the group just to troll. The moderators put a lot of effort into keeping their groups closed from potential troublemakers who might disrupt discussion and prevent other members from participating. In many groups, member lists were curated so that moderators could judge each applicant before giving approval. Sometimes, they would discuss the merits of acceptance with their fellow moderators. Judgement of suitability for the group was passed by inspecting information on an applicant's profile page and, in particular, their liked pages and other group memberships. Membership of some strictly moderated Facebook groups were perceived as a recommendation and proof of an applicant's good behavior. As the following quote shows, some moderators were adept at detecting potential trolls and troublemakers by looking for certain signs:

> They can be very discreet. Once there was someone who had created numerous troll accounts and each account had the same background picture. But you don't see that until you put the profiles side by side to compare them. It's a dog whistle for the like-minded; an invitation to troll. When they see those certain signs in the profile, they will join in the trolling. (Interviewee 10)

The groups being studied were at different stages in their life cycle, which in turn affected their member approval policy. Some groups were rather new and at a growth stage with plenty of applications for membership coming in, and so moderators would accept new members daily. In these groups, the moderators did not scrutinize applications as carefully because they wanted new members to join. Moreover, some groups were at a stage of saturation, and the moderators were satisfied with current membership levels and user activity. In these groups, they were a little reluctant to accept new members and stated that they did not want the group to grow any bigger because new members would bring increased workload and the potential of trouble.

For mature and well-functioning groups, newcomers pose a greater risk as they might challenge existing norms and express their disagreement. In this way, they need extra moderation and guidance. They may be perceived as a threat to the power and authority

of established members, particularly in situations when many newcomers are accepted at the same time (Honeycutt, 2005). However, for online communities to sustain themselves and grow, new members must still be occasionally accepted. One of the groups in this study was at a terminal stage with diminishing user activity and rare new applicants. In this group, moderation policy was very strict and only a few newcomers were accepted. Eventually, this led to the group diminishing.

Similar findings about boundary control have been made in prior research of online communities. In particular, elite members control access to a community and monitor who is allowed to participate in conversations (Honeycutt, 2005; Weber, 2011). If a newcomer fails to conform to group norms, the elite refuse to accept that member into the group unless they admit their ignorance of the norms (Weber, 2011). However, disruption caused by newcomers can be useful for a community as it helps moderators and members to identify and define rules and boundaries. Moderators screen members because they want to minimize damage and avoid additional work. As a moderator from a well-functioning and stable group said: "I admit that when I judge someone's suitability as a member of the group, I think about the potential workload. If their profile information gives the impression that this is a quarrelsome person, I might not approve them" (Interviewee 3).

When members are carefully screened and their views are seen to be similar to the group consensus, moderators are more likely to apply softer moderation strategies. One way of conceptualizing moderation strategies is to divide these into soft and hard, based on how much moderation restricts users' activities in the group. Personal discussions, in which a moderator contacts the member privately and notifies them about questionable behavior, were mentioned as the softest form of moderation, whereas excluding a member from the group either temporarily or permanently was generally considered to be the hardest form of moderation. The strength of moderation can also be defined based on its visibility to users. In this sense, screening members in advance is a soft form of moderation: When someone is not accepted, this does not leave any trace for group members to see because outsiders are not allowed to post in the group. However, declining someone's right to participate can be considered the strongest limitation that a moderator can apply to users. Warning someone discretely in person and hiding someone from discussions without their knowledge are invisible forms of moderation, whereas public interventions in a discussion, bans, and removals are usually visible to the whole group, and therefore might harm one's reputation within the group. Private discussion between moderator and member was considered a discrete form of moderation because it is invisible to other members and allows the person in question to save their face in the group. Private discussion was used particularly in situations when a troublemaking user was well known in the group, or when moderators suspected that a user might regret their behavior afterwards, for example due to drunkenness. As one moderator said: "In the moderation business, I often feel that we need to protect people from themselves, like preventing them for causing harm to themselves" (Interviewee 6).

Many moderators reported using hard forms of moderation, namely bans and removals, actively and without previous negotiation. However, doing so is likely to cause unwanted reactions in the group as removing someone can create tension and criticism among group members. Because hard moderation is visible to other members, banning or kicking someone out of the group publicly is likely to give even more visibility to their opinions. Removing members tends to invoke critical discussion about censorship and sympathy for the removed person. Moderators admitted that trolls sometimes use this approach to cause extra fuss and reaction from others, and therefore they intentionally provoke moderators with the aim of being punished publicly. For this reason, moderators need to be careful in how they deal with provocative content:

> Some people just want to get to say that "the moderation sucks in this group, I am leaving now." And then others begin to wonder if there is something wrong with this group. We have to remove the "I'm leaving" notes because they are used with the intention of harming the group and the good spirit between people. (Interviewee 10)

The findings point out that in successful groups, moderators have a strong sense of community and belonging to their group, which is an important factor behind them actively volunteering for moderation work. Moderators with strong feelings of belonging to their group feel ownership and are committed to taking care of and nurturing the community by continually monitoring content and membership. If founders and moderators do not feel any ownership or obligation to look after their group, it is more likely to become filled with arguments and misinformation. Some moderators mentioned cautionary examples of abandoned, non-moderated groups that attracted political actors who used them to spread their own political agendas. Eventually, the group would drift away from its original purpose.

4.2. Power Dynamics Between Moderators and Members

Another main aim of this study is to understand the power dynamic between moderators and group members, and to find out whether some members' views are given more priority than others. In the interviews, moderators were asked if they perceive comments from all group members as being equal in terms of their value and contribution to the group. They uniformly stated that some social media users are better in having their opinions heard and accepted by the group, while others'

opinions remain ignored. They also described the key characteristics of an influential group member, saying that such a person's individual skills are the most important reason for salience. Asked what most important qualities are for being taken seriously by others, the moderators emphasized good writing and argumentation skills and sound knowledge and expertise of matters under discussion. They also stressed that merely being vocal and active in the group does not make someone influential. If a participant is not good at expressing their opinions in written form or lacks grammatical skills, it is harder for them to be perceived as credible in an online discussion. In addition to skills, being famous through offline activity and having a strong reputation based on previous history as a member also contribute to a member's salience. It seems that a member's personal friendship with the moderator is not perceived as a factor that leads to greater visibility and salience; but instead, these active key members tend to develop closer relationships with moderators and gain more influence over time and activity in the group. These key members have an important role in directing discussion as their opinions are more valued and trusted than those of less-known members. Some moderators admitted that it is difficult to moderate these salient members, and as a result, they are given more freedom to express their views.

Becoming a prominent member who is valued in the group leads to a virtuous circle. Those who are active and comment regularly become well-known among their peers and gain more prestige over time. Eventually, anything they say is likely to receive positive attention. However, moderators admitted that salience could sometimes become harmful for the group if prominent members dominate discussions and draw all the attention to themselves, while others do not receive attention and feedback to their comments. One moderator said that she would encourage less visible members by liking and commenting on their posts:

> When someone famous in the group posts something, she gets loads of likes, whereas someone who's not so good at expressing her ideas and does not have the same status receives no reaction. I try to be on the side of the underdog and comment with something positive like "yeah, that's great" just to show some empathy. (Interviewee 6)

Contrary to traditional media, the success of social media sites relies on users' activity. Even though moderators possess a considerable amount of power over discussion within the group, its members are not completely powerless, and they can influence the course the group takes through their participation. In the interviews, the moderators mentioned a couple of ways how moderated members or users who had been removed would resist their moderation policy. The first is by flagging and reporting content to Facebook's own moderation team. Sometimes, when content receives a number of flags, it will be removed, even if it is not against Facebook own policy. Flagging content is thus used to bypass local group moderators and question their power (see also Gillespie, 2018). The second way is to create a competing Facebook group that would discuss the same topics and be intended for the same audience, albeit with a different moderation policy that is tailored to addressing the perceived faults of the original group. Among the groups studied, there were some examples of groups that had been created in protest to some other group's moderation policy as their founders had felt they had been treated unfairly.

Eventually, users have the power to keep communities alive by participating or abandoning them, which makes social media groups highly dependent on their membership, and particularly on those who are active contributors and valued by moderators and their peer members. Users can abandon groups if they are not satisfied, and without users creating and updating content, groups will eventually die. This demonstrates how in the context of social media, the power relationship between gatekeepers and the gated remains dynamic and can change.

5. Discussion and Conclusions

As exposure to news, opinions, and political information increasingly occurs through social media, scholars have expressed their concern about its narrowing and polarizing effect on the information that people are exposed to. Research has confirmed social media users' tendency to network with those who have similar opinions, which is identified as a main driving force behind polarization (Boutyline & Willer, 2017; Lewis et al., 2011). These communities of like-minded people are suspected of amplifying individuals' existing beliefs and restricting the free flow of information, which is harmful for the formation of balanced political views, and thus for deliberative democracy. In connection with this scholarly discussion, which often takes place in relation to algorithmic moderation, this study shows how information is filtered in politically motivated grassroots groups by human moderators. Opacity of moderation policy, which has been named as the main problem in the way that platforms conduct moderation (Gillespie, 2018; Roberts, 2016), is also present in moderation done by volunteers. If users are unaware of the filtering that is performed on their behalf, they do not know what information is left out and why, which leads to their participation in public debate being inadequate or even biased.

Social media platforms are major gatekeepers of information as the selection of content is inherent to them. Moderated private groups, such as those presented in this study, provide in many ways a prominent base for the polarization of views, especially if they do not allow dissenting opinions. Controlling access is an effective way to maintain the homogeneity of a group, and hence, moderation may pose a risk that groups eventually develop into echo chambers. Similar to the study

by Kalsnes and Ihlebæk (2021), this study views moderation practices that are concealed from group members as problematic because they are an obstacle for deliberative democracy and personal development. When users are not accepted into a group to begin with, they are moderated and silenced even before they have participated. This study proposes that transparency throughout all decisions and strategies of moderation is important for civic discussion.

The present study has some limitations as the findings rely solely on interviews with moderators. In order to analyze the moderation process as a whole, and better understand power dynamics between gatekeepers and the gated, future research needs to include viewpoints from all stakeholder groups involved, and particularly from those who are the object of moderation.

This study shows that the work of volunteer moderators encompasses a much wider range of activities than simply hiding or removing content, which are named as the main elements of Facebook's approach to moderation (Kalsnes & Ihlebæk, 2021). Controlling access by curating member lists is a major part of moderation in groups on Facebook; however, it has often been left unexplored in prior related studies. Through continuous boundary control, moderators define the group's ideals for those who are inside and outside of group, as well as for themselves. By focusing on their groups' boundaries, the moderators in this study were shown to view outsiders as the biggest threat to the groups and norms. When boundaries are blurred, the existence of the group may be threatened and open to attack. Access to the group is regulated in order to not only maintain group norms and cohesion of views but also to avoid harder forms of moderation. Hard moderation—namely by restricting users' participation or altering it by removing or editing content—occurs in all groups, but because these activities can affect harmony and bring consequences, moderators would rather prevent such incidents by carefully screening potential members. In particular, when the group is in a state that is satisfying for moderators and key members, accepting new members may pose risks.

Volunteer moderators have a challenging task of responding to members' expectations while maintaining the group's main purpose through their everyday moderation tasks. Prior studies have suggested that user-driven communities tend to develop non-democratic structures, so that some users tend to gain more privileges and visibility than others (Keegan & Gergle, 2010; Shaw & Hill, 2014). This study recognizes the existence of "elite members" who have more visibility and power in relation to moderators and get their messages across better than others. Usually, these active members are viewed as beneficial for online communities as these are dependent on their contributions (e.g., Malinen, 2015), but sometimes they can be harmful for the group and its dynamics. A salient member can draw attention to themselves, and so discourage others from contributing.

In the current high-choice social media environment, information transmission has become more direct, but many of the mechanisms through which information flows from producers to users still remain invisible. This study has revealed how volunteer moderators control political discussion in groups on Facebook, and its findings show how they hold a disproportionate amount of power over group members. The relationship between gatekeepers and the gated is thus asymmetrical and unilateral; gatekeepers possess a range of tools for limiting, or even preventing, the participatory opportunities of the gated (see also Dovbysh, 2021). This article has approached gatekeeping in social media as means of control that involves several controlling practices that moderators can use towards group members. Focusing particularly on one of these strategies—boundary control—the findings show how it is used as a discrete but effective way of controlling content. It prevents unwanted users from participating, but at the same time, it is not evident to group members.

Acknowledgments

My warmest thanks to the moderators who participated in this study and shared their experiences. I would also like to thank the anonymous reviewers and the editors of this thematic issue for their useful comments and encouragement in revising this article. This study is part of the author's post-doctoral project NETGATE (2019–2022) funded by Academy of Finland, decision number 321608.

Conflict of Interests

The author declares no conflict of interests.

References

Barzilai-Nahon, K. (2008). Toward a theory of network gatekeeping: A framework for exploring information control. *Journal of the American Society for Information Science and Technology*, *59*(9), 1493–1512.

Bennett, W. L., & Segerberg, A. (2012). The logic of connective action. *Information, Communication & Society*, *15*, 739–768.

Bessi, A., Petroni, F., Del Vicario, M., Zollo, F., Anagnostopoulos, A., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016). Homophily and polarization in the age of misinformation. *The European Physical Journal Special Topics*, *225*(10), 2047–2059.

Boutyline, A., & Willer, R. (2017). The social structure of political echo chambers: Variation in ideological homophily in online metworks. *Political Psychology*, *38*(3), 551–569.

Choak, C. (2012). Asking questions: Interviews and evaluations. In S. Bradford & F. Cullen (Eds.), *Research and research methods for youth practitioners* (pp. 90–112). Routledge.

Creswell, J. W. (2013). Qualitative inquiry and research design: Choosing among five traditions. SAGE.

Del Vicario, M., Bessi, A., Zollo, F., Petronic, F., Scala, A., Caldarellia, G., Stanleye, H. E., & Quattrociocchia, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, *113*, 554–559.

Dovbysh, O. (2021). New gatekeepers in town: How groups in social networking sites influence information flows in Russia's provinces. *Social Media + Society*, *7*(2). https://doi.org/10.1177/20563051211013253

Dutton, W. H. (2009). The fifth estate emerging through the network of networks. *Prometheus*, *27*(1), 1–15.

Fiesler, C., Feuston, J. L., & Bruckman, A. S. (2015). Understanding copyright law in online creative communities. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (pp. 116–129). ACM.

Gachau, J. N. (2016). The role of social media in participatory democracy: A case study of Facebook groups. *Proceedings of the 19th International Conference on Supporting Group Work* (pp. 467–472). ACM.

Gagliardone, I., Gal, D., Alves, T., & Martinez, G. (2015). *Countering online hate speech*. UNESCO.

Gillespie, T. (2018). *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.

Gorwa, R. (2019). What is platform governance? *Information, Communication & Society*, *22*(6), 854–871.

Gustafsson, N. (2012). The subtle nature of Facebook politics: Swedish social network site users and political participation. *New Media & Society*, *14*(7), 1111–1127.

Honeycutt, C. (2005). Hazing as a process of boundary maintenance in an online community. *Journal of Computer-Mediated Communication*, *10*(2). https://doi.org/10.1111/j.1083-6101.2005.tb00240.x

Jhaver, S., Appling, D. S., Gilbert, E., & Bruckman, A. (2019). "Did you suspect the post would be removed?" Understanding user reactions to content removals on Reddit. *Proceedings of the ACM on human-computer interaction*, *3*(CSCW), 1–33.

Kalsnes, B., & Ihlebæk, K. A. (2021). Hiding hate speech: Political moderation on Facebook. *Media, Culture & Society*, *43*(2), 326–342.

Keegan, B., & Gergle, D. (2010). Egalitarians at the gate: One-sided gatekeeping practices in social media. *Proceedings of the 2010 ACM conference on Computer Supported Cooperative Work* (pp. 131–134). ACM.

Kushin, M. J., & Kitchener, K. (2009). Getting political on social network sites: Exploring online political discourse on Facebook. *First Monday*, *14*(11). https://doi.org/10.5210/fm.v14i11.2645

Langlois, G., Elmer, G., McKelvey, F., & Devereaux, Z. (2009). Networked publics: The double articulation of code and politics on Facebook. *Canadian Journal of Communication*, *34*(3). https://doi.org/10.22230/cjc.2009v34n3a2114

Lewis, K., Gonzalez, M., & Kaufman, J. (2011). Social selection and peer influence in an online social network. *Proceedings of the National Academy of Sciences*, *109*(1), 68–72.

Malinen, S. (2015). Understanding user participation in online communities: A systematic literature review of empirical studies. *Computers in Human Behavior*, *46*, 228–238.

Matias, J. N. (2019). The civic labor of volunteer moderators online. *Social Media + Society*, *5*(2). https://doi.org/10.1177/2056305119836778

Meraz, S., & Papacharissi, Z. (2013). Networked gatekeeping and networked framing on #Egypt. *The International Journal of Press/Politics*, *18*(2), 138–166.

Myers West, S. (2017). Raging against the machine: Network gatekeeping and collective action on social media platforms. *Media and Communication*, *5*(3), 28–36.

Pariser, E. (2011). *The filter bubble: What the internet is hiding from you*. Penguin Press.

Park, N., Kee, K. F., & Valenzuela, S. (2009). Being immersed in social networking environment: Facebook groups, uses and gratifications, and social outcomes. *Cyberpsychology & behavior*, *12*(6), 729–733.

Pruchniewska, U. (2019). "A group that's just women for women": Feminist affordances of private Facebook groups for professionals. *New media & society*, *21*(6), 1362–1379.

Roberts, S. T. (2016). *Commercial content moderation: Digital laborers' dirty work*. https://ir.lib.uwo.ca/commpub/12

Ruckenstein, M., & Turunen, L. L. M. (2020). Re-humanizing the platform: Content moderators and the logic of care. *New media & society*, *22*(6), 1026–1042.

Seering, J., Wang, T., Yoon, J., & Kaufman, G. (2019). Moderator engagement and community development in the age of algorithms. *New Media & Society*, *21*(7), 1417–1443.

Shaw, A., & Hill, B. M. (2014). Laboratories of oligarchy? How the iron law extends to peer production. *Journal of Communication*, *64*, 215–238.

Sormanen, N., & Dutton, W. H. (2015). The role of social media in societal change: Cases in Finland of fifth estate activity on Facebook. *Social Media + Society*, *1*(2). https://doi.org/10.1177/2056305115612782

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*(3), 755–769.

Terranova, T. (2000). Free labor: Producing culture for the digital economy. *Social Text*, *18*, 33–58.

Warren, A. M., Sulaiman, A., & Jaafar, N. I. (2014). Facebook: The enabler of online civic engagement for activists. *Computers in Human Behavior*, *32*, 284–289.

Weber, H. L. (2011). Missed cues: How disputes can socialize virtual newcomers. *Language@ Internet*, *8*(5), 1–18.

Zhu, Q., Skoric, M., & Shen, F. (2017). I shield myself from thee: Selective avoidance on social media during political protests. *Political Communication*, *34*(1), 112–131.

About the Author

Sanna Malinen is postdoctoral researcher in economic sociology at the University of Turku, Finland. She has studied online communities for over a decade, focusing particularly on users' collective information production and their dynamic roles as information providers and consumers. Her ongoing postdoctoral project examines how information is selected, controlled, and circulated in online public sphere, and how these practices shape power dynamics between gatekeepers and the gated.

COGITATIO

Media and Communication (ISSN: 2183–2439)
2021, Volume 9, Issue 4, Pages 82–92
https://doi.org/10.17645/mac.v9i4.4227

Article

The Agency of Journalists in Competitive Authoritarian Regimes: The Case of Ukraine During Yanukovich's Presidency

Esther Somfalvy * and Heiko Pleines

Department of Politics and Economics, Research Centre for East European Studies at the University of Bremen, Germany; E-Mails: somfalvy@uni-bremen.de (E.S.), pleines@uni-bremen.de (H.P.)

* Corresponding author

Submitted: 27 February 2021 | Accepted: 14 May 2021 | Published: 21 October 2021

Abstract
On the example of Ukraine during the Yanukovich presidency (2010–2014) this article explores which factors support journalists' agency in relation to censorship pressure in a competitive authoritarian regime. It shows that a critical mass of journalists existed who reacted to censorship pressure with rejection. Based, first of all, on 31 semi-structured interviews, we examine the working conditions of prominent national journalists and analyse how they describe their role and motivations. We argue that the nature of competitive authoritarianism offers journalists opportunities for critical reporting, but that it is individual characteristics of journalists—including professional ethics, networks, and job mobility—which define whether and how the respective opportunities are used.

Keywords
authoritarian regime; censorship; competitive authoritarianism; journalism; Ukraine; Yanukovich

Issue
This article is part of the issue "Media Control Revisited: Challenges, Bottom-Up Resistance, and Agency in the Digital Age" edited by Olga Dovbysh (University of Helsinki, Finland) and Esther Somfalvy (Research Centre for East European Studies at the University of Bremen, Germany).

1. Introduction

Scholarly discussions of media control in non-democratic regimes often focus on the macro perspective, investigating ways of media control, like ownership or legal regulation of media. What is less studied is the agency of journalists in this context that enables them to react differently to the resulting pressure. The literature dealing with democracies offers some suggestions on possible influences that are rooted in individual journalists' profiles. According to Helmueller and Mellado (2015), role perception affects news content created by journalists, whereby role conceptions vary more between countries than within them (van Dalen et al., 2012), highlighting the importance of the country-specific context. However, in-depth studies of journalistic reaction to political pressure in a non-democratic context are largely missing beyond recent studies on self-censorship (Schimpfössl et al., 2020).

That is why we are interested in which conditions foster the agency of journalists when faced with political pressure in a competitive authoritarian regime. In this kind of regime, governments attempt to maintain the image of a functioning democracy while using a multitude of tools, including the systematic manipulation of institutions, to tilt the political "playing field" in their favour (Levitsky & Way, 2010, p. 1). In the case of media, this leads to practices of censorship that are covert and heterogenous (Levitsky & Way, 2010, pp. 8–9). Business people take over major mass media as part of their deals with ruling political elites; in this case, censorship pressure is exerted by media owners: a phenomenon which has been described as "media capture" (Mungiu-Pippidi, 2008; on Ukraine, see Ryabinska, 2017, pp. 59–69). In the

case of Ukraine, for example, since the early 2000s, about two-thirds of television viewers have been watching news programmes from stations that are owned by the country's most influential business people, so-called oligarchs (Pleines, 2016, p. 124).

This analysis looks at Ukraine from 2010 to 2014, a period during which the country is considered an exemplary case of a competitive authoritarian regime (Levitsky & Way, 2010; Pleines, 2012). The media system in Ukraine has been described as being characterized by "poverty, small size of the market, strong politicization and control of media, and low professionalism" (Dobek-Ostrowska, 2015, p. 35). In this situation, the TV market is not driven by market logic, but rather seen as means through which to accumulate political influence (Ryabinska, 2011, p. 5). Media also depend on political advertising. For-pay political advertising in the form of so-called *dzhynsa* (hidden for-pay advertising) was common during the 2005–2010 election campaigns (Grynko, 2012, p. 263; Ryabinska, 2011, p. 10). These developments had affected reporting, e.g., leading to news channels ignoring facts deemed "inconvenient" for the government and an imbalance in reporting on the government compared to the opposition (Ryabinska, 2017, pp. 71–72).

Critical independent media faced selective pressure from the ruling political elites and related business networks. One example of an outlet under pressure was the television channel TVi, which lost its general broadcasting licence already in 2010, the first year of the Yanukovich presidency. In 2012 it was investigated by the tax authorities and lost access to major satellite TV distributors. Finally, in April 2013 it saw an unfriendly takeover ("Independent TV station under increasing threat," 2012). At the same time, this example demonstrates the agency of journalists. Journalists do not have to acquiesce to mounting pressure on their work by the state or oligarchic media owners. In the face of the TVi takeover, they went on strike. Later some moved on to found the independent TV channel Espreso.tv. Another prominent case was *Forbes Ukraine*, which had conducted an intensive investigation into state procurement systematically favouring companies allegedly close to president Yanukovich. The journal was taken over by a businessman with very close links to Yanukovich and the investigative journalists left (Tuchynska, 2013).

In sum, the conditions under which journalists in Ukraine worked were characterized by pressures on their reporting levied upon them by the political environment, including most prominently oligarchs.

2. Reacting to Censorship Pressure on Media Under Autocracy

Much has been written about the fact that (mostly non-democratic) governments put pressure on media and individual journalists to influence reporting (Akhrarkhodjaeva, 2017; McMillan & Zoido, 2004). Political pressure for self-censorship is often created in violation of journalists' civil rights using violence, selective law enforcement, and manipulated charges. It is often assumed that in anticipation of potential troubles most journalists in authoritarian regimes engage in self-censorship, meaning that they shy away from reporting on certain individuals or topics due to their anticipation of the consequences their reporting might otherwise have. In this sense, self-censorship is "often understood in relationship to censorship," as Schimpfössl et al. (2020, pp. 1–11) summarize the debate.

Pressures regarding what to report on and how to report may result in very different actions by those they target. Numerous studies on a variety of countries show that many journalists react to strong and violent pressure with wide-ranging self-censorship (Kenny & Gross, 2008; Nadadur, 2007; Tong, 2009; Yesil, 2014). Many journalists also develop conformism and claim that they report in an appropriate way (Schimpfössl & Yablokov, 2014). Journalists may also choose to respond with ethical justifications about why engaging in some forms of self-censorship is appropriate (Skjerdal, 2008).

Scholarly analyses of the phenomenon are complicated by the fact that pressures are often communicated euphemistically, unlike, for example, the often cited *temnyki* (from *temnii*- "dark," directives from the presidential administration to media) of the Kuchma era in Ukraine, which provided daily instructions to journalists what (not) to cover in reporting (Grynko, 2012, p. 263). In cases where explicit censorship is absent, the "red lines" that should not be crossed when reporting are subject to interpretation and change (Fedirko, 2020; Zeveleva, 2020) so that journalists might fail to adequately anticipate the consequences of their actions. Under such conditions, journalistic self-censorship is seen as the result of the "interplay between free will, coercion and obligation" (Fedirko, 2020, p. 13).

Accordingly, the result of pressures on journalist may be something entirely different than self-censorship since journalists have agency in choosing from a large repertoire of responses to such pressures. Some journalists, while explicitly accepting some self-censorship pressure, test its limits (Lee & Chan, 2008). Others reject such pressure outright and get into conflict with chief editors and media owners, continuing to write stories about topics that may be perceived as problematic. Furthermore, some journalists chose to resist pressure by quitting the job in protest, participating in protests, or funding independent news outlets.

What, then, enables certain journalists to resist censorship pressure and continue with critical reporting (and risk open conflict with owners or politicians)? As there is no universal framework that could be applied to the case, this study is exploratory. However, the literature on what influences media content can give some hints as to which areas to examine (Shoemaker & Reese, 1996).

It is safe to expect that the political regime plays an important role. Self-censorship exists in democracies

(Kohut, 2000). However, it is unique to authoritarian regimes that the *content* subjected to public self-censorship pressure relates, first of all, to the performance of the political regime and its representatives, thus addressing a core issue of political media reporting. There is some literature in a political science tradition that examines the conditions under which journalists in authoritarian regimes have more or less leeway to shape media contents. The most prominent explaining factor is the overall degree of authoritarian control (Stier, 2015). Indeed, in authoritarian regimes, the pressure towards self-censorship is usually coordinated country wide. As a result, a specific bias in reporting is not restricted to an individual media outlet, but applied across major media, thus considerably reducing media pluralism.

However, it is important to note that in a *competitive* authoritarian regime the censorship pressure on media can strongly vary, as it is not centrally administered by a state agency. Accordingly, at the level of media organizations, we can expect a difference in pressure by media type, where TV is supposed to be most strongly controlled and the internet to offer most opportunities for pluralism (Heinrich & Pleines, 2018, p. 5). Furthermore, we know from the Russian case that pressure from owners, transmitted via loyal editors, affects content (Fredheim, 2017) and differs in line with the position of the owners vis-à-vis the political regime.

Moreover, there are individual-level factors that influence reporting. A recent study finds that organizational factors have the largest influence on news production, while individual predispositions matter far less (for the summary of the literature see Hanitzsch et al., 2019, p. 105). However, as we are interested in more than media content production (where editorial decisions matter more), factors relating to the own professional path, personal network, or the question of how individual journalists make sense of the context they work in are as important. In this context, how journalists reflect upon their conditions and their role perception has found to have an effect on news content (Helmueller & Mellado, 2015).

3. Research Design

Our analysis aims to answer the question of what enables journalists to resist to pressures of self-censorship in a competitive authoritarian regime. As the related state of research is limited, we provide an exploratory case study examining the role of journalists working at major national media in Ukraine during the Yanukovich presidency (2010–2014). Ukraine at this time is not only a typical case of a competitive authoritarian regime, but it is also marked by a developed media system and a larger number of renowned investigative journalists. The period under study allows us to examine a consolidating competitive authoritarian regime and its crisis in the wake of the Euromaidan protests from November 2013 to February 2014.

Our exploratory case study is first of all based on 31 semi-structured interviews with Ukrainian journalists who were active during the period under study. A reputational sampling method was used to identify interviewees. With the reputational approach, we follow the strand of the literature on elite interviewing which argues for the identification of the most relevant interview partners (those with the respective reputation) instead of opting for a probability sample. Tansey (2007, p. 765), for example, states that in such cases the aim "is to obtain information about well-defined and specific events and processes, and the most appropriate sampling procedures are thus those that identify the key political actors—those who have had the most involvement with the processes of interest." To identify the key actors of relevance for a specific study, a snowball technique is often used so that the reputational approach has also been described as a different kind of snowball sampling (Farquharson, 2005).

The study presented here combines desktop research and the snowball technique to identify those prominent national journalists who had rejected censorship pressure during the Yanukovich presidency. Accordingly, our analysis describes the situation of leading national journalists, while the experience of local journalists in smaller cities was likely different. It is, thus, important to note that while Ukraine stands for many similar competitive authoritarian regimes with a media market with considerable oligarch influence, the individual journalists selected for the study do not constitute a cross-section of the profession. Our selection focuses on (often prominent) cases of independent journalists reacting with insubordination to censorship pressures and cross-references them with experiences of journalists from other groups, thus examining the conditions that enabled some to resist censorship pressure that was present throughout the system.

Questions addressed during the interviews were inspired by the literature and covered the following areas: employment history, job motivation, and journalistic neutrality. Specifically relating to their employment situation during the Yanukovich presidency, questions addressed hiring practices, work routines, the relationship between editor (owners/managers) and journalist, topics covered or difficult to cover, and self-censorship. Where applicable interviewees were asked about their experiences after an ownership change, quitting a job, or establishing a new media outlet during the period under study.

The interviews were conducted by Esther Somfalvy (Research Centre for East European Studies at the University of Bremen [FSO]) and by partners of the research project Comparing Protest Actions in Soviet and Post-Soviet Spaces, which is organized by Heiko Pleines. These partners were the Public Sociology Laboratory (PS-Lab) and the Foundation for the Preservation of the History of Maidan (FPHM), which for the interviews cooperated with the Ukrainian Institute of National Memory

(UINP; see Kovtunovych & Pryvalko, 2015). Interviews were conducted in English, Russian, or Ukrainian depending on the preference of the respondent. With a few exceptions, all interview partners agreed to the use of their full names in quotations. As they have a long experience of seeing their names in print (all are journalists, some are persons of public interest), they know what this agreement means. Some have explicitly objected to anonymized statements. A list of interview partners is provided in Table 1, a full documentation of the interview process has been published on the DiscussData Platform (Somfalvy & Pleines, 2021; on DiscussData see Heinrich et al., 2019).

As the interviews were conducted after the end of the Yanukovich presidency, they are likely to contain some hindsight bias. However, as we are focusing on journalists rejecting censorship pressure, i.e., being critical of Yanukovich while he was in power, this is less of an issue. That there is no strong political bias in our interview sample is also confirmed by the fact that several respondents commented critically on the state of media freedom in the post-Maidan period. In the interviews, specific examples were usually given to substantiate more general statements. Moreover, the position of these journalists was at the time demonstrated by their public actions, e.g., writing open letters or quitting their job in protest. Still, whenever possible, we have triangulated interview statements by talking to several people from the same media outlet and by checking additional documents and media reporting at the time.

Table 1. List of interview partners.

Interviewee	Interview by	Organization 2013	Status 2013	Date of interview
Atanasov, Vitalii	FSO	*Fokus*	Regular author	2019-05-14
Babich, Bogdana	FPHM	Spilno.tv	Co-founder	2014-12-09
Berdynskykh, Kristina	FPHM	*Korrespondent*	Journalist, blogger	2015-04-30
Burdyga, Igor	FSO	*Kommersant/Vesti Reporter*	Senior business correspondent	2019-05-14
Davidenko, Boris	FSO	*Forbes*	Editor-in-chief	2019-05-13
Gumenyuk, Nataliya	FSO	Several channels	Freelancer	2019-05-15
Ivanchenko, Roman	FPHM	Interfax	Journalist	2016-10-04
Kalnysh, Valerii	FSO	*Kommersant*	Editor-in-chief	2019-05-17
Kapshuchenko, Yulia	UINP	n.a.	Journalist	2014-10-21
Kapustin, Andrey	UINP	Freelancer	Journalist, blogger	2015-06-04
Karagyaur, Vladimir	FPHM	Spilno.tv	Volunteer	2015-04-07
Khardy, Mar'yana	UINP	Freelancer	Photojournalist	2014-03-21
Melykh, Olga	FSO	*Ukrainian Week*	Communication officer	2019-05-15
Nerodyk, Inna	UINP	Channel 5	Journalist	2014-11-06
Paskhover, Aleksandr	FSO	*Korrespondent*	Editor business	2019-05-14
Petrenko, Galina	FSO	*Marketing Media Review*	Editor-in-chief	2019-05-14
Piddubyi, Oleksandr	UINP	Freelancer	Journalist	2014-08-22
Portnikov, Vitalii	FSO	Espreso.tv	Editor-in-chief	2019-05-17
Romaniuk, Roman	FSO	UNIAN	Editor news	2019-05-17
Rybak, Vitalii	FSO	Local newspaper in Vynitsa	Freelancer	2019-05-16
Samofalov, Andrii	FSO	*Forbes Ukraine*	Journalist	2019-05-13
Sadomtseva, Galina	FPHM	Spilno.tv	Volunteer	2015-07-13
Shara, Anatolii	FSO	Maidan media, *Tyzhden*	Freelancer	2019-05-15
Shirochenko, Vladimir	FPHM	Freelancer	Photojournalist	2015-04-03
Shovkoshitnyi, Radion	FPHM	TV channel "Business"	Journalist	2015-05-02
Sokolenko, Natal'ya	UINP	*Centr-UA*	Journalist	2015-01-14
Yasenchuk, Aleksandr	UINP	Local media in Chernihiv	Journalist	2014-03-20
Zaklets'kii, Oleksandr	FPHM	Freelancer	Photojournalist	2014-11-28
Anonymous (No. 16)	PS-Lab	n.a.	Journalist	2014-07-10
Anonymous (No. 31)	PS-Lab	n.a.	Editor	2014-07-XX
Anonymous (No. 48)	PS-Lab	n.a.	Journalist, blogger	2014-07-17

A further source of information for this study is a survey on journalistic ethics and professionalism conducted in May 2013 among 52 journalists (of which 30 are from Kiev) working for print, radio, television, and internet media (Ilko Kucheriv Democratic Initiative Foundation, 2013). For a closer analysis of the context, media reports on issues relating to the media environment, media freedom and censorship have been included in the case, as well as other relevant documents, like legal acts, statements issued by agencies, and experts.

4. Rejecting Self-Censorship Pressure in Ukraine (2010–2014)

In order to explain the rejection of self-censorship pressure by a number of prominent national journalists in Ukraine during the Yanukovich presidency, we first present their perception of this pressure. We then go on to describe three key features of the media landscape which are associated with higher levels of journalistic agency and offered opportunities for critical reporting: namely the existence of niches, professional ethics, and a flexible job market. The period of the Euromaidan protests in 2013–2014, which was accompanied by increasing pressure on journalists, will be analysed in a separate section.

4.1. Perception of Censorship Pressure

The interviews support the expectation that in competitive authoritarian regimes, like Ukraine under President Yanukovich, political censorship pressure is not applied equally throughout the country and across all media. Regarding the creation of news content, there is a broad spectrum: Some respondents say they only ever experienced conflict between correspondents and editors that were part of daily reporting and had no qualms to fight for their point of view (Interview Burdyga) and could decline to report on any topic they were asked to cover (Interview Anonymous No. 16). That the exact "red lines" for reporting were not always clear is supported by the account of a journalist at *Kommersant Ukraine* who after an ownership change checked the fresh copies for changes made to articles and was surprised to find that none had been made (Interview Burdyga).

The majority of journalists say that they did expect some pressure from owners. Oligarchic ownership, according to Boris Davidenko, former editor-in-chief at *Forbes Ukraine*, often led to the creation of a list of people and topics that were not to be covered in a negative way (Interview Davidenko). Everybody in the industry, interviewees claim, knew that certain people were better left alone (Interview Romaniuk—at the time news editor at UNIAN news agency). Pressures to censor topics were sometimes quite explicit. A journalist explained that once the medium she worked at was sold, the editorial policy changed quickly and openly, as journalists "were told that the investigations related to Yanukovich, his family and team were impossible" (Interview Berdinskykh). Davidenko recalls that when *Forbes* was taken over by a businessman close to President Yanukovich, journalists were promised higher salaries if they stopped critical reporting and simultaneously threatened with dismissal if they did not (Interview Davidenko). At UNIAN news agency journalists were fined if they negatively covered President Yanukovich, as was made public by five former employees in an open letter in October 2012 ("Barometr svobodi slova," 2012).

While media owners often intervened systematically in the reporting of journalists, in the perception of journalists, state organs only reacted when red lines were crossed. However, writing something that was deemed "unacceptable" resulted in threats (Interview Shara). Another journalist reported having been personally followed and spied upon by secret services in 2012 while working at TVi (Interview Portnikov). These personal accounts are supported by figures from a report compiled by Reporters Without Borders ("Moment of truth," 2012) that finds systematic state violence against journalists. This violence later escalated further in the wake of the Euromaidan protests.

4.2. Agency Based on Position, Knowing About "Niches"

Differences in censorship pressure felt across the profession can partly be explained by the nature of the outlet and by different preferences of the owners that are—to some extent—known to journalists. One respondent described self-censorship at TV stations as "some kind of Stockholm syndrome," where it became the standard way of functioning and was no longer challenged (Interview Paskhover). There was less pressure in low-profile print publications. One editor believed that the perceived unimportance of the publication within the owner's portfolio made it possible to enjoy some freedom in their reporting (Interview Petrenko). Another interview partner reports that the pressure on the staff of Radio Vesti grew once its popularity grew (Interview Kalnysh). The interviews also support the perception that online media had more freedom. According to an interviewee:

> I worked on the newspaper's website… we were a little bit disconnected, we put online news that was not much related to the newspaper—we allowed ourselves to be objective there. Journalists who worked directly on the newspaper—it was harder for them. (Interview Anonymous No. 31)

Interview partners also report on the existence of niches by topic or type of reporting. For example, an interview partner claims that investigative reporting could be done relatively unhindered (personal communication), and another one reported that business media was such a niche (Interview Burdyga). The claim about business media being a niche where reporting could be done

relatively freely requires some context, given the structure of the Ukrainian media market with its oligarchic owners. The factor most often mentioned to explain differences in censorship pressure is, in fact, the position of the owner. This is supported by a poll of 52 journalists conducted in May 2013 in which 38 named business people with political interests as one of the biggest obstacles to press freedom, while only 27 chose ruling politicians (Ilko Kucheriv Democratic Initiative Foundation, 2013). Moreover, as owners have different interests, censorship pressure does not lead to uniform media reporting (for which in Ukraine the term "oligarchic pluralism" has been coined; Ligachova, 2015).

Another example of how ownership affects what topics are considered to be unproblematic is provided by an interview partner who recounts his experiences at a newspaper outside the capital, which was controlled by the local authorities. This meant that the actions of the national government could be criticized, while local matters had to be reported about very carefully (Interview Rybak).

Finally, it must be noted that the perceived interests of the owner and the consequences for their work are subject to journalists' interpretations. It may not even always be clear to journalists who owns their outlet (Interview Burdyga). Opaque ownership structures make it more difficult to factor in negative reactions to potentially contentious actions.

4.3. Motivation and Professional Ethics

Professionalism and a certain understanding of how a journalist should report—"objectively" or "neutrally"—play an important part in the interviewees' role understanding. Several of them also express the belief that they would quit a job rather than compromise their integrity.

When the journalists interviewed, who were chosen for their prominent rejection of censorship pressure during the Yanukovich presidency, describe their outlook on their job, it becomes apparent that although their motivation for becoming a journalist may vary—and is often related either to chance or to a need to earn some money—at some point during their career idealism or the idea of having a societal role to play prevailed. One such perspective-changing event is described by Shara, who witnessed violence by the police during the protests on Maidan and describes it as a "bifurcation point" (Interview Shara; see also Budivska & Orlova, 2017, p. 147).

There are different ideas about what it means to be a journalist and how "objective" or "neutral" reporting can be. However, all of our respondents highlight that journalists should follow professional standards. With that, they position themselves against any censorship pressures which they perceive as unethical. Although they use different terminology (referring to objectivity and neutrality) their ideas resemble Hanretty's (2011) concept of journalistic autonomy. It is also important to note that our respondents identify with different versions of "neutral" reporting, not with support for specific political camps or ideologies. It is also indicative that—although they all opposed the Yanukovich regime—many of them are also highly critical of the media freedom record of the post-Maidan government.

Some journalists self-describe as particularly driven by ideals, a fact which was known to colleagues and gave them special status (Interview Anonymous No. 31). Another journalist claims that it was clear to the new owners of her outlet that she would not write anything as a favour (Interview Berdinskykh).

As a consequence of this attitude, when the own outlet will no longer accommodate their professional standards, these journalists will quit their job and move elsewhere. During our research for this study, we have counted 13 such cases involving over 100 journalists for the period under study (Somfalvy & Pleines, 2021). One journalist reports already having quit a well-paid job for one where he could actively oppose Kuchma in 2004 (Interview Shyrochenko). Many journalists quit without having a clear idea of where to go next (Interview Berdinskykh).

4.4. Job Market

The idealistic interest in working as a journalist as well as the need for some income means that the important question from the perspective of the journalists who do not fit into the pro-regime media is whether there are alternative job opportunities available, as this makes the risk of leaving or the threat of being fired less dramatic. To what extent this is the case in competitive authoritarian regimes has so far not been examined systematically. Accordingly, it is an important insight into the functioning of censorship pressures that most of our respondents who lost or quit their job during the Yanukovich presidency did not experience problems finding a new one. The Euromaidan of 2013–2014 is also reported as having provided young journalists with a window of opportunity in the job market (Interview Melykh).

Collectively, our respondents point to three explaining factors for alternative job opportunities for critical journalists within the media system:

(1) Relatively large demand for journalists in a situation of rather unattractive employment conditions. The financial situation of the media allows for little financial support for regular high-quality reporting, although some funding for investigative journalism is available (Interview Gumenyuk). Many do not stay in the profession long-term (Interview Davidenko; Interview Petrenko). Low salaries and bad working conditions also mean that journalists do not have much to lose when quitting their job.
(2) Personal networks. Recruitment for new positions is often based on pre-existing personal networks

and recommendations. Several interview partners found jobs through friends (Interview Ivanchenko), or left their jobs together with team members and helped one another to find new employment (Interview Burdyga). Some journalists report choosing a place of work because they had friends already working there (Interview Shara). Editors may also feel responsible for their colleagues (Interview Kalnysh). In this context, it is also important to note that in Ukraine under Yanukovich seemingly there was no blacklist of journalists critical to the regime. When one of our respondents, an editor-in-chief, lost his job at a pro-Yanukovich newspaper in a conflict about reporting, he got a phone call with a job offer on the very same day, and only after he had started to work at the new media outlet did he realize that it was owned by a pro-Yanukovich oligarch, causing him to quit again (Interview Kalnysh).

(3) Opportunities to set up new media outlets. A larger number of journalists who lost or quit their job got involved in the creation of new media outlets. When *Korrespondent* was taken over in 2013, a whole team of journalists quit and later went on to found *Novoe Vremya*. Portnikov reports breaking with TVi due to their refusal to report on the Euromaidan protests (Interview Portnikov). He set up Espreso.tv with a team largely made up of former TVi employees in November 2013. Similarly, Bogdana Babich and her team set up Spilno.tv which provided live broadcasts from the protests. Other journalist-driven projects were Hromadske.TV and Channel 112. That this was possible is due to the nature of competitive authoritarianism which formally guarantees media freedom, and factually only restricts it when it challenges the ruling elites, which small start-up media at least initially did not do.

As discussed in the section on the research design this study focuses on prominent journalists working for national media. Clearly, the opportunities outlined above are not available to all journalists. Most importantly, the media landscape outside the capital is usually much less varied, so that job mobility may simply not be feasible due to the lack of job offers. Moreover, regional media were in many cases under stricter and more unified control either by the regional state administration or one dominating oligarch than national ones.

However, even for prominent journalists this relative job mobility only exists as long as there are niche publications accommodating journalist who offend the regime, and an official blacklist of journalists does not exist. This started to change during the later stage of the Yanukovich presidency, as we will discuss in the following section.

5. The Euromaidan Protests

The beginning of the protests in late November 2013 on the Maidan Nezalezhnosti, a central square in Kiev, in response to the announcement that the government would not, as anticipated, sign an association agreement with the European Union (which is why the protests were later referred to as "Euromaidan") took place during a time of growing concern for the future of independent media.

Ukrainian journalists in 2013 felt a tightening grip of Yanukovich's associates on media. Journalists recall that in 2013 the space for independent journalism was rapidly shrinking, as a "big shopping" of media assets was going on among oligarchs (Interview Kapustin). This prompted some to fear that if nothing changed, soon there would be no space for independent reporting left, and some journalists were already prepared to switch profession or leave the country (Interview Anonymous No. 48). Paskhover, business editor at *Korrespondent* magazine at that time, recalls a meeting in November 2013 where together with a colleague they announced that they wanted to quit over the new editorial policy:

> When we said that we would leave, we were honestly told that there would be nowhere to go soon. They were going to buy everything. It was the first time I thought about changing my profession. I just really had the feeling that, in general, there would be nowhere to go professionally. Well, and then two weeks later Maidan happened.

A substantial number of journalists was present at and participating in the protests from the start. This was partially because the initial call for protest was a post on Facebook by Mustafa Nayyem, a journalist who addressed his friends and colleagues. Journalists mentioned as the reason for initially going to Maidan their ties to the organizers (Interview Babich; Interview Kapustin), or "sociological curiosity" (Interview Berdinskykh) about the unfolding events. As the protests gained momentum, more journalist recognized their magnitude and expressed a feeling that this was simply the place they had to be (Interview Khardy; Interview Piddubyi). Professional networks seem to play a role in whether journalists went to the protests at the early stage. A network of like-minded colleagues already existed through their involvement in the Stop Censorship! movement founded in 2010 (Budivska & Orlova, 2017, p. 140).

The protests received little attention from state-owned media at the beginning and were covered mostly by small independent media (including online TV channels that became popular due to this coverage). Initially, there was relative freedom to report on the events, which is often attributed to the fact that the oligarchs owning the majority of the media did not take a side, which left editors and journalist to decide for themselves

how to cover the events and whether to participate in the protests (for a detailed account of how reporting changed throughout the events see Ryabinska, 2015; Szostek, 2014). When protests grew in size, journalists working at some pro-Yanukovich media were not allowed to cover them, but apart from that were left to their own devices. A journalist working at a TV channel linked to Yanukovich, the channel Business, recounts his experience:

> In the middle of the Maidan [protests], we were broadcasting "Swan Lake" non-stop. We'd come to work, but we didn't do the news. On the one hand, it's good that you're not on vacation and you get paid. At this time, instead, you could have gone to Maidan in peace. On the other hand, you should show up for work, but you're not working. (Interview Shovkoshytnyi; similarly Interview Anonymous No. 31).

The situation for freelancers was different. While they could simply decide to go to Maidan (Interview Pidduby), freelancers could have difficulty proving they were there on an assignment and could be prosecuted for participation in a riot. The consequence, as one interviewee explains, was the blurring of the line between those participating as citizens, as journalists, or as combatants (Interview Zaklets'kii).

The Euromaidan protests brought increasing repression, including physical violence, some specifically targeting journalists and opposition media ("Raids on three opposition media," 2013). Another catalytic moment came in January 2014, when the parliament passed repressive media legislation. Galina Petrenko who during that time was the editor-in-chief of a marketing publication recounts that she and her colleagues were discussing how to react—either by publishing an issue containing only white pages or by writing about the issue. The decision to write was taken with a feeling of "while we can still talk, well, until the law comes into force, we have to talk. Our job is to talk. Well, that's why we talked as much as we could" (Interview Petrenko).

Regardless of whether they started attending the protests in a journalistic capacity or as activists, many interviewees say that they attempted to keep these roles separate. They did this by separating between day-job covering the events and activism in their spare time (Interview Ivanchenko). In many instances, this proved impossible as the events unfolded—just as suggested by the literature that finds a blurring of boundaries between journalism and activism (Ligachova, 2015; Szostek, 2014). For example, one journalist also explains that only later did she and her colleagues realize that they were already engaged in activism beyond their purely journalistic work (Interview Petrenko).

6. Conclusion

The research presented in this article examines what enabled a group of Ukrainian journalists to reject censorship pressure and exercise agency over their reporting. The case study demonstrates that a critical mass of journalists existed under competitive authoritarianism in Ukraine who rejected censorship pressure.

It can be stated that the dispersed control over media assets that is typical for competitive authoritarian regimes, and which in the case of Ukraine is exercised by oligarchs, is an important element of what could be described as opportunities for critical journalism. The diverse strategies of media owners—with additional differentiation by media type and visibility—can explain why journalists are exposed to different levels of self-censorship pressure, which means that niches for critical reporting exist. A second important element of the opportunities for critical journalism is the high degree of job mobility, often based on professional networks. Journalists were also sharing the perception that they knew about the rules and niches where they could report according to their standards, switching jobs if they felt their professional ethics compromised. Importantly, related job mobility was not hampered by any coordinated blacklisting of critical journalists. A third element of the opportunities for critical reporting is the possibility to register and organise independent media outlets. All of these elements facilitated the agency of journalists under the competitive authoritarian regime established by Yanukovich.

The opportunities are thus dependent on having some media pluralism in place. The perception that this pluralism was being threatened by the Yanukovich regime was also shared by many interviewees. This may have contributed to mobilization as the Euromaidan protests started in late 2013, as journalists were facing the question of how to position themselves vis-à-vis the regime. Russia in the last two decades serves an illustrative case of what happens when a regime obtains control over a large proportion of the media sphere, which by itself is a sign of the regime becoming fully authoritarian (Pleines, 2020; Somfalvy, 2020).

The journalists who used the opportunities for critical journalism were all driven by professional ethics focusing on journalistic autonomy. That this was not about empty words was proven by several mass resignations of entire journalistic teams in the years 2012–2013. Such collective action also points to another important factor explaining critical journalism—the embeddedness in networks of like-minded colleagues. This was relevant not only for job security but also for joint projects like the foundation of new media outlets and joint activism.

These findings on journalists under authoritarianism resonate with the literature on other professions' agency, which is linked to their motivation for the rejection of authoritarian control and collective action. For example, studies on lawyers suggest that professional ethics and encountering the violation of rights might foster mobilization against an authoritarian regime (Kazun & Yakovlev, 2017). On the other hand, lawyers rely on state structures to a higher degree than

journalists, who might also be more mobile. This suggests that journalists could, by nature of their profession, have more agency than lawyers when working under competitive authoritarianism. Hence, a comparison with how other professions resist authoritarianism could foster a broader understanding of how individuals function within and make use of opportunity structures their profession provides based on how it relates to the wider regime's context.

Acknowledgments

This article has been produced as part of the research project Media Control as Source of Political Power: The Role of Oligarchs in Electoral Authoritarian Regimes, which is conducted by the Research Centre for East European Studies at the University of Bremen and receives financial support from the Deutsche Forschungsgemeinschaft (DFG)—grant No. 391270526. Additional interviews included in the analysis presented here have been conducted as part of the research project Comparing Protest Actions in Soviet and Post-Soviet Spaces, which is organized by Heiko Pleines at the Research Centre for East European Studies at the University of Bremen with financial support from the Volkswagen Foundation.

Conflict of Interests

The authors declare no conflict of interests.

References

Akhrarkhodjaeva, N. (2017). *The instrumentalisation of mass media in electoral authoritarian regimes: Evidence from Russia's presidential election campaigns of 2000 and 2008*. Ibidem.

Barometr svobodi slova za zhovten 2012 roku [Freedom of speech barometer for October 2012]. (2012, November 7). *Institute for Mass Information*. https://imi.org.ua/monitorings/barometr-svobodi-slova-za-jovten-2012-roku

Budivska, H., & Orlova, D. (2017). Between professionalism and activism: Ukrainian journalism after the Euromaidan. *Kyiv-Mohyla Law and Politics Journal*, *3*, 137–156. https://doi.org/10.18523/kmlpj120120.2017-3.137-156

Dobek-Ostrowska, B. (2015). 25 years after communism: Four models of media and politics in Central and Eastern Europe. In B. Dobek-Ostrowska & M. Glowacki (Eds.), *Democracy and media in Central and Eastern Europe 25 years on* (pp. 11–46). Peter Lang.

Farquharson, K. (2005). A different kind of snowball: Identifying key policymakers. *International Journal of Social Research Methodology*, *8*(4), 345–353. https://doi.org/10.1080/1364557042000203116

Fedirko, T. (2020). Self-censorships in Ukraine: Distinguishing between the silences of television journalism. *European Journal of Communication*, *35*(1), 12–28. https://doi.org/10.1177/0267323119897424

Fredheim, R. (2017). The loyal editor effect: Russian online journalism after independence. *Post-Soviet Affairs*, *33*(1), 34–48. https://doi.org/10.1080/1060586X.2016.1200797

Grynko, A. (2012). Ukrainian journalists' perceptions of unethical practices: Codes and everyday ethics. *Central European Journal of Communication*, *2*(5), 259–274.

Hanitzsch, T., Hanusch, T., Ramaprasad, J., & de Beer, A. S. (Eds.). (2019). *Worlds of journalism: Journalistic cultures around the globe.* Columbia University Press.

Hanretty, C. (2011). *Public broadcasting and political interference*. Routledge.

Heinrich, A., Herrmann, F., & Pleines, H. (2019). Transparency and quality assessment of research data in post-Soviet area studies: The potential of an interactive online platform. *Journal of Eurasian Studies*, *10*(2), 136–146. https://doi.org/10.1177/1879366519850698

Heinrich, A., & Pleines, H. (2018). The meaning of "limited pluralism" in media reporting under authoritarian rule. *Politics and Governance*, *6*(2), 103–111. https://doi.org/10.17645/pag.v6i2.1238

Helmueller, L., & Mellado, C. (2015). Professional roles and news construction: A media sociology conceptualization of journalists' role conception and performance. *Communication and Society*, *28*(3), 1–11.

Ilko Kucheriv Democratic Initiative Foundation. (2013, June 5). *Svoboda slova v Ukraini: Zagal'nonacional'ne i ekspertne opitovannia* [Freedom of speech in Ukraine: A nation-wide expert survey]. https://dif.org.ua/en/article/svoboda-slova-v-ukraini-zagalnonatsionalne-y-ekspertne-opituvannya

Independent TV station under increasing threat. (2012, September 7). *Reporters Without Borders*. https://rsf.org/en/news/independent-tv-station-under-increasing-threat

Kazun, A., & Yakovlev, A. (2017). Who demands collective action in an imperfect institutional environment? A case study of the profession of advocates in Russia. *Journal of Eurasian Studies*, *8*(1), 60–71. https://doi.org/10.1016/j.euras.2016.08.001

Kenny, T., & Gross, P. (2008). Journalism in Central Asia: A victim of politics, economics, and widespread self-censorship. *The International Journal of Press/Politics*, *13*(4), 515–525. https://doi.org/10.1177/1940161208324644

Kohut, A. (2000). Self-censorship: Counting the ways. *Columbia Journalism Review*, *39*(1), 42-43.

Kovtunovych, T., & Pryvalko, T. (2015). *Maidan vid pershiyi osoby: 45 Istorii Revolyuciyi hidnosti* [First-person Maidan: 45 stories of the Revolution of Dignity]. Ukrainian Institute of National Memory.

Lee, F., & Chan, J. (2008). Organizational production of self-censorship in the Hong Kong media. *The International Journal of Press/Politics*, *14*(1), 112–133.

https://doi.org/10.1177/1940161208326598
Levitsky, S., & Way, L. A. (2010). *Competitive authoritarianism: Hybrid regimes after the Cold War.* Cambridge University Press.
Ligachova, N. (2015, April 20). Media oboz ili media avangard? [Media bandwagon or media avant-garde]? *MediaSapiens*. http://osvita.mediasapiens.ua/trends/1411978127/mediaoboz_ili_mediaavangard
McMillan, J., & Zoido, P. (2004). How to subvert democracy: Montesins in Peru. *Journal of Economic Perspectives*, *18*(4), 69–92.
Moment of truth for freedom of information, concern on eve of elections. (2012, October 24). *Reporters Without Borders*. https://rsf.org/en/news/moment-truth-freedom-information-concern-eve-elections
Mungiu-Pippidi, A. (2008). How media and politics shape each other in the new Europe. In K. Jakubowicz & M. Sukosd (Eds.), *Finding the right place on the map: Central and Eastern European media change in a global perspective* (pp. 87–100). Intellect.
Nadadur, R. D. (2007). Self-censorship in the Pakistani print media. *South Asian Survey*, *14*(1), 45–63. https://doi.org/10.1177/097152310701400105
Pleines, H. (2012). From competitive authoritarianism to defective democracy. Political regimes in Ukraine before and after the Orange Revolution. In S. Stewart, M. Klein, A. Schmitz, & H. Schröder (Eds.), *Presidents, oligarchs, and bureaucrats: Forms of rule in the post-Soviet space* (pp. 125–138). Ashgate.
Pleines, H. (2016). Oligarchs and politics in Ukraine. *Demokratizatsiya*, *24*(1), 105–127.
Pleines, H. (2020). Media control as source of political power: Differentiating reach and impact. *Russian Analytical Digest*, *258*, 2–7.
Raids on three opposition media as pro-European protests dispersed. (2013, December 9). *Reporters Without Borders*. https://rsf.org/en/news/raids-three-opposition-media-pro-european-protests-dispersed
Ryabinska, N. (2011). The media market and media ownership in post-communist Ukraine. *Problems of Post-Communism*, *58*(6), 3–20. https://doi.org/10.2753/PPC1075-8216580601
Ryabinska, N. (2015). Ukraine: Local media on the Euromaidan protests. *Cultures of History Forum*, *2015*. https://doi.org/10.25626/0031
Ryabinska, N. (2017). *Ukraine's post-communist mass media: Between capture and commercialization*. Ibidem.
Schimpfössl, E., & Yablokov, I. (2014). Coercion or conformism? Censorship and self-censorship among Russian media personalities and reporters in the 2010s. *Demokratizatsiya*, *22*(2), 295–311.
Schimpfössl, E., Yablokov, I., Zeveleva, O., Fedirko, T., & Bajomi-Lazar, P. (2020). Self-censorship narrated: Journalism in Central and Eastern Europe. *European Journal of Communication*, *35*(1), 3–11. https://doi.org/10.1177/0267323119897801
Shoemaker, P., & Reese, S. (1996). *Mediating the message: Theories of influence on mass media content*. Longman.
Skjerdal, T. S. (2008). Self-censorship among news journalists in the Ethiopian state media. *African Communication Research*, *1*(2), 185–206.
Somfalvy, E. (2020). Shrinking niches for independent journalism: The case of Vedomosti. *Russian Analytical Digest*, *258*, 8–11.
Somfalvy, E., & Pleines, H. (2021). *Data collection for Journalists in competitive authoritarian regimes. The case of Ukraine 2010–2014, v. 1.0* [Data set]. Discuss Data. https://discuss-data.net/dataset/9211b92e-a806-4e6c-ad75-d031a910b9f1
Stier, S. (2015). Democracy, autocracy, and the news: The impact of regime type on media freedom. *Democratization*, *22*(7), 1273–1295. https://doi.org/10.1080/13510347.2014.964643
Szostek, J. (2014). The media battles of Ukraine's Euromaidan. *Digital Icons*, *2014*(11), 1–19. http://www.digitalicons.org/issue11/the-media-battles-of-ukraines-euromaidan
Tansey, O. (2007). Process tracing and elite interviewing: A case for non-probability sampling. *PS: Political Science & Politics*, *40*(4), 765–772.
Tong, J. (2009). Press self-censorship in China: A case study in the transformation of discourse. *Discourse & Society*, *20*(5), 593–612. https://doi.org/10.1177/0957926509106412
Tuchynska, S. (2013, August 8). Journalists to leave Forbes in protest over new owner. *Kyiv Post*. https://www.kyivpost.com/article/content/ukraine-politics/journalists-to-leave-forbes-in-protest-over-new-owner-327912.html
van Dalen, A., de Vreese, C., & Albæk, E. (2012). Different roles, different content? A four-country comparison of the role conceptions and reporting style of political journalists. *Journalism*, *13*(7), 903—922. https://doi.org/10.1177/1464884911431538
Yesil, B. (2014). Press censorship in Turkey: Networks of state power, commercial pressures, and self-censorship. *Communication, Culture, and Critique*, *7*(2), 154–173. https://doi.org/10.1111/cccr.12049
Zeveleva, O. (2020). Towards a bourdieusian sociology of self-censorship: What we can learn from journalists adapting to rapid political change in Crimea after 2014. *European Journal of Communication*, *35*(1), 46–59. https://doi.org/10.1177/0267323119897798

About the Authors

Esther Somfalvy is a research fellow at the at the Research Centre for East European Studies at the University of Bremen. Her research interests include comparative authoritarianism studies, political institutions (parliaments, elections), and media.

Heiko Pleines is head of the Department of Politics and Economics, Research Centre for East European Studies at the University of Bremen and professor of comparative politics at the University of Bremen. His research focuses on the functioning of authoritarian regimes in the post-Soviet region, including prominently the role of mass media.

COGITATIO

Media and Communication (ISSN: 2183–2439)
2021, Volume 9, Issue 4, Pages 93–103
https://doi.org/10.17645/mac.v9i4.4204

Article

Resisting Perceived Interference in Journalistic Autonomy: The Study of Public Service Media in Slovakia

Marína Urbániková

Department of Media Studies and Journalism, Masaryk University, Czech Republic; E-Mail: urbaniko@fss.muni.cz

Submitted: 18 February 2021 | Accepted: 12 April 2021 | Published: 21 October 2021

Abstract
Autonomy is of paramount importance for journalism, but there is little empirically based knowledge of how journalists cope when it is threatened. Using a case study approach, this contribution examines a newsroom conflict that took place in the public service Radio and Television of Slovakia. It started when the new director general, a person believed to have ties to one of the coalition political parties, was elected by the parliament in 2017, and it culminated in layoffs and resignations of more than 30 reporters and editors in 2018. The case study is based on semi-structured interviews (N = 16) with the journalists who decided to quit in protest of what they called "creeping political pressure," those whose contracts were not prolonged, those who decided to stay at their jobs, and the members of the previous and the new management. Building on the interviews and document analysis, the article inductively develops a classification scheme for resistance practices the journalists used to cope with the perceived interference with their professional autonomy that came from within their media organisation. These practices include having internal discussions, voicing concerns during newsroom meetings, writing an internal letter to the management, meeting with the management, establishing a trade union, requesting mediation, writing an open letter to the viewers and listeners, publicly criticising the management in the media, voluntarily asking to be re-assigned to another topic area or position in order to avoid interference, staying at one's job in open opposition to the management, and resigning in protest.

Keywords
autonomy; interference; newsroom conflict; pressure; public service media; resistance practices; RTVS; Slovakia

Issue
This article is part of the issue "Media Control Revisited: Challenges, Bottom-Up Resistance and Agency in the Digital Age" edited by Olga Dovbysh (University of Helsinki, Finland) and Esther Somfalvy (University of Bremen, Germany).

1. Introduction

Professional autonomy is one of the holy grails of journalism. The freedom of journalists to make and follow their own norms and rules of practice is one of the key ideal-typical values upon which journalism's ideology is based (Deuze, 2005). It is what makes journalism a profession (Freidson, 1994) or, in Bourdieu's terms, what makes it a separate field (Bourdieu & Wacquant, 1992). Given its paramount significance, important questions arise: How do journalists react when they feel that their autonomy is threatened? What options and measures do they have to handle what they perceive as undue interference?

This study sheds more light on the different ways that professional autonomy can be defended in practice. More concretely, it focuses on the resistance practices used to deal with perceived internal pressure from within the news organisation (i.e., from management). It employs a qualitative case study approach (Yin, 2018) to analyse a newsroom conflict that took place in the Slovak public service broadcaster Radio and Television of Slovakia (RTVS) in 2017 and 2018. The conflict started when the new director general, a person believed to have ties to one of the then-coalition political parties, was installed by the parliament, and it culminated with the layoffs and resignations of more than 30 reporters and editors who complained of "creeping political pressure" (Jančáriková, 2018).

Even though the importance of professional autonomy in journalism is well documented in the scholarly

literature, resistance practices used by journalists when their autonomy is in jeopardy have rarely been studied (as pointed out, e.g., by Barrios & Miller, 2020). Empirical studies mostly focus on the perceived level of journalistic autonomy in various countries (Ahva et al., 2017; Hughes et al., 2017), or, from the opposite angle, on exploring various types and forms of interference in journalistic autonomy (Akhrarkhodjaeva, 2017b; Goyanes & Rodríguez-Castro, 2019) and on the extent of the journalists' experience with this interference (Clark & Grech, 2017; Hiltunen, 2019). However, the question of how journalists actually deal with the pressure and interference is less often addressed, and if so, available studies have focused mostly on external political interference that occurs in flawed democracies and authoritarian or hybrid regimes (Ataman & Çoban, 2019; Barrios & Miller, 2020; Slavtcheva-Petkova, 2019). Another relevant stream of literature, the research on conflicts in public service media, zeroes in on cultural clashes between the content makers and top managers who are responsible for administering and running "the factory" (Nissen, 2014), and on concrete cases when the independence of public service broadcasters was breached and journalistic autonomy constrained (Dragomir, 2017; Dzięciołowski, 2017; Koivunen, 2017). Again, an emphasis on the array of possible resistance practices of dissatisfied journalists is missing. This is where this study steps in.

The case of the newsroom conflict in the public service medium in Slovakia is of interest for three reasons. First, as already suggested, previous research has focused primarily on external political interference in non-democratic countries, unlike this study, which examines the case of (perceived) internal interference from within the media organisation in a European Union country with a relatively high ranking for democracy and press freedom. Slovakia is currently ranked 42nd out of 167 countries in the 2019 Democracy Index (The Economist Intelligence Unit, 2020), and 33rd out of 180 countries in the 2020 World Press Freedom Index (Reporters Without Borders, 2020). Second, this study contributes to our knowledge of journalism culture in Central and Eastern Europe, a region which has, compared to Western societies, been studied considerably less (e.g., Standaert et al., 2019). As journalistic autonomy certainly did not belong to the core ideal-typical values of journalism during the communist times, the question arises whether and to what extent the journalists working for the public service broadcaster in Slovakia adopted autonomy as an essential value that was worth defending. And third, journalistic autonomy is of particular importance in the realm of public service media because their *raison d'être* is independence from both political and economic pressure. Therefore, although journalists in general are expected to defend their autonomy when they feel it is under attack, this expectation is reasonably higher in the case of journalists working for public service media, which makes RTVS an interesting case to study.

This article is organised as follows: It first reviews the literature on the resistance practices that journalists use to cope with interference in their autonomy and, drawing from organisational studies, reviews the literature on the practices that employees use to express their dissent (Section 2); it describes the research method and data (Section 3); it analyses the newsroom conflict and introduces an inductively developed classification scheme of resistance practices through 16 semi-structured interviews with the journalists and managers from the RTVS newsroom (Section 4); Section 5 is the summary and conclusion.

2. Literature Review: Resisting Interference and Voicing Dissatisfaction

2.1. Journalistic Autonomy and Coping With its Encroachment

Journalistic autonomy is the "latitude that a practitioner has in carrying out his or her occupational duties" (Weaver et al., 2007, p. 70). It refers to "the extent to which journalists can make decisions free of pressures from management, commercial factors, as well as other forces that reside inside the news environment" (Reich & Hanitzsch, 2013, p. 135). Journalistic autonomy can be threatened by the interference of various actors, either internally, from within the journalistic field (e.g., editors, managers), or externally, most notably from the political or economic fields. Interference can be described as threats or inducements which cause or attempt to cause journalists to act in a particular fashion (Hanretty, 2011, p. 5), or as methods used to "influence journalists with the objective of shaping editorial content" (Hiltunen, 2019, p. 5). Thus, interference does not only refer to direct interventions (or attempts at interventions) in the journalistic content, but also to all sorts of pressure to discipline the journalists and make them act according to the interests of the sources of this pressure. In the case of internal actors (most notably editors), in order to qualify their actions as interference, these actions must be journalistically unwarranted.

Available studies that explore the practices used by journalists when they feel their autonomy is endangered focus mostly on cases of external political interference that occur in flawed democracies and authoritarian or hybrid regimes (see Ataman & Çoban, 2019, for Turkey; Barrios & Miller, 2020, for Columbia; Fedirko, 2020, for Ukraine; Slavtcheva-Petkova, 2019, for Russia). These practices include: seeking support from editors (Barrios & Miller, 2020; Slavtcheva-Petkova, 2019); sharing or handing over sensitive stories to other colleagues or working anonymously (Ataman & Çoban, 2019; Barrios & Miller, 2020); using social media to attract followers in order to raise the costs for the potential sources of the pressure (Barrios & Miller, 2020); trying to solve the problems by directly contacting the sources of the pressure (Slavtcheva-Petkova, 2019); practicing borderless

and cross-border journalism (Ataman & Çoban, 2019); and using the support of international actors (Ataman & Çoban, 2019; Barrios & Miller, 2020).

Studies that focus on democratic countries mostly examine how journalists deal with commercial interference, either internal or external (see Borden, 2000, for the U.S.; Goyanes & Rodríguez-Castro, 2019, for Spain; Hanusch et al., 2017, for Australia and Germany). Regarding external commercial interference, coping practices include avoiding negative accounts about a product or service (Goyanes & Rodríguez-Castro, 2019; Hanusch et al., 2017); not reporting about a product or service at all (Hanusch et al., 2017); and being more careful about meeting journalistic norms in sensitive cases (Goyanes & Rodríguez-Castro, 2019). Regarding internal commercial pressure, according to Borden (2000), journalists use open protest, sabotage (e.g., making decisions without consulting higher levels of management), principled compromise (i.e., concession in order to accomplish basic journalistic goals), and "trump cards," which suggest that non-compliance with standard journalism would lead to a loss of credibility.

Although these studies provide useful insights into how journalists cope with interference in different contexts, none of them examine a situation similar to this case study: A case of perceived internal interference that comes from within the media organisation and that is interpreted by dissenting journalists as the lack of professional skills on the side of the management (in the best case) and as politically motivated (in the worst case). Thus, this case study explores a unique situation where journalists perceive the interference of their superiors as journalistically unwarranted with only speculation for the underlying motivations.

2.2. Employee Dissatisfaction in an Organisation: Exit and Voice in Journalism

Besides research that focuses on how journalists deal with perceived encroachment on their autonomy, organisational studies are the second relevant stream of scholarship upon which this contribution is based. Specifically, it builds up the literature on the practices that employees use to express their dissent for the organisation where they work. In his widely cited work, Hirschman (1970) summarised that people respond to the decline in the performance of organisations mainly by exit or voice. They either leave the organisation (or stop buying its products), or they voice their discontent and exert pressure upon the organisation to improve its performance. Loyalty is the key moderating variable: Less loyal employees and customers are more likely to exit; more loyal ones are more likely to use their voice (Hirschman, 1970). Later research identified a variety of other predictors such as: organisational commitment; job satisfaction; perceived justice, trust, and fairness; psychological contract violation; job alternatives; employability; and psychological safety (Aravopoulou et al., 2017; Subhakaran & Dyaram, 2018).

Farrell (1983) enriched the exit-voice conceptualisation with two other options. Besides quitting or voicing their discontent, dissatisfied employees can opt for loyalty (i.e., stick with the organisation and patiently wait for improvement) or neglect (i.e., passive behaviour, withdrawal, and hostility). Focusing on employees' voice strategies, Gorden (1988) distinguished four types that are based on active-passive and constructive-destructive dimensions. Accordingly, as summarized by Kassing (2002), employees' voice can be active constructive (e.g., making suggestions, argument, union bargaining), passive constructive (e.g., listening, quiet support, unobtrusive compliance), active destructive (e.g., complaining, ingratiation, verbal aggression, antagonistic exit), or passive destructive (e.g., murmurings, apathy, withdrawal).

In the field of journalism and media studies, the exit-voice-loyalty-neglect model has been applied for several purposes: to explore and evaluate journalist resistance to business constraints (Borden, 2000); to analyse the responses of journalists to ethical dilemmas (Saldaña et al., 2016); or to examine career choices of journalists (Akhrarkhodjaeva, 2017a; Davidson & Meyers, 2016). Arguably, media organisations have several specific features that need to be taken into account when applying the exit-voice-loyalty-neglect model to journalists' behaviour. These traits make journalists' position when it comes to expressing dissent different from that of, for example, assembly line workers. Journalism is a specific profession with a high potential for conflict and employee dissatisfaction. The very nature of journalism—the construction of media representations of reality—offers much room for ideological tensions. Also, journalism can be considered as a semi-profession (Tunstall, 1971) without universally accepted standards and rules of practice and without a clear definition of what good professional performance is. This can be another source of potential disputes between journalists and management.

Journalism entails not only a high risk of conflicts but also specific requirements for their resolution. When journalists disagree with their employer's editorial policy, the stakes are high, and so are the societal demands placed on them. It is in journalists' vital interest to defend their autonomy as, without it, they lose their authority and can no longer be considered professionals (Borden, 2000). However, at the same time, journalists who would openly voice their discontent face extremely high costs (Borden, 2000). In some cases, using voice may mean quitting with "the possible prospects of not being able to return to the profession until conditions change" (Akhrarkhodjaeva, 2017a, p. 8). Also, many journalists hold the view that those who disagree with the editorial policy of their media organisation are free to quit and change to another one (Schimpfossl & Yablokov, 2014). In short, it seems that in journalism, there is an increased pressure on journalists to opt for voice or exit as opposed to loyalty or neglect—and of course, that comes with a price.

3. Data and Method

This article aims to explore the resistance practices used by the journalists working for RTVS to cope with what they perceived as the undue interference of their superiors upon their professional autonomy during a newsroom conflict that took place in 2017 and 2018. The study asks not only what resistance practices the journalists used, but also what was their order, and how was the selection of individual practices related to the broader journalists' strategies of how to respond to the unsatisfactory conditions in the newsroom. Even though the conflict affected both the radio and television divisions of RTVS, this study focuses on the television only because the clash was more dramatic, more closely followed by the public, and led to more staff resignations.

As a broader methodological strategy to learn more about the journalistic resistance practices, a case study approach—"an empirical method that investigates a contemporary phenomenon (the 'case') in-depth and within its real-world context" (Yin, 2018, p. 15)—was employed. Conflicts where journalists publicly complain of interference with their autonomy serve as a useful research opportunity because they allow for the capture of the variability of individual and collective resistance practices, as well as the sequences. To explore the case of RTVS, I conducted 16 semi-structured interviews with the main actors of the conflict (e.g., journalists, managers) and supplemented it with document analysis (e.g., news articles, an open letter written by journalists, the management's response) as a form of triangulation.

As pointed out by Nissen (2016), conflicting parties in public service media organisations typically offer different interpretations of what happened. What one stakeholder describes as a brave defence of autonomy and independence, another interprets as a consequence of unsatisfactory performance and a lack of loyalty. Obviously, semi-structured interviews cannot reveal which side of the conflict in RTVS was "right" or "wrong," nor do they prove whether or to what extent the new RTVS management interfered with the professional autonomy of the journalists. However, this methodological approach can shed more light on how the journalists reacted once they perceived the interference. To verify (to the extent possible) the basic facts, I checked for inconsistencies in the versions and interpretations of conflicting parties and compared them with descriptions of individual events as captured in the news and other documents. All the interviewees agreed on the elementary explanation of what happened, and the document analysis also supported these findings. As expected, what was not agreed upon were individual actors' motivations and the interpretation of key values that needed to be protected.

Given that the subject of the analysis is an organizational conflict, it is essential to mention the RTVS' internal structure. RTVS was created in 2011 following a merger of Slovak Television with Slovak Radio, and the director general supervises both the radio and television divisions. They are elected (and potentially also dismissed) by a simple majority of votes in the parliament. The term of office is for five years, and the same person may be elected to only two consecutive terms. When it comes to news making, the highest-ranking role (right below the director general) is the head of television and radio news and current affairs. Their direct subordinates are the editor-in-chief of television news and current affairs and the editor-in-chief of radio news and current affairs. One level below the editors-in-chief are the team leaders who lead the rank-and-file reporters.

The selection of conversational partners was led by an effort to cover the key groups of actors and to maximise the diversity in the sample from the viewpoint of position, gender, age, and length of working experience. Participants were selected based on the author's knowledge of the case. With one exception, none of the addressed participants declined the invitation to participate in the research study. The conversational partners fall into five key groups: the journalists whose contracts were not prolonged by the new management (1 participant); the journalists who resigned in protest (4); the journalists who decided to stay at their jobs (5; however, one of them resigned shortly after the interview); newly appointed managers (4); and members of the previous management who resigned (2). Managers are defined as people who oversaw editors and rank-and-file reporters (e.g., team leaders, the head of television and radio news and current affairs). The years of experience of the 5 female and 11 male participants ranged from 3 years to more than 20. In the following text, for the sake of brevity, I use the umbrella term "journalists" to denote the dissenting part of the newsroom that included a major part, though not all, of the reporters and editors.

The interviews were carried out between July 2018 and September 2019. All of the interviews were recorded, anonymised, transcribed verbatim, and subjected to coding in Atlas.ti. To ensure better anonymity for the conversational partners, the generic feminine pronoun is used throughout the text when referring to the participants. To analyse the data, I used thematic analysis, "a method for identifying, analysing and reporting patterns (themes) within data" (Braun & Clarke, 2006, p. 79). The coding and analysis process followed the analytic procedure suggested by Braun and Clarke (2006): It started with familiarising with the data and generating initial codes, continued with searching for themes (collating the codes in potential themes) and reviewing themes (including the creation of a thematic map), and ended with defining and naming themes, and producing the report.

4. Analysis: The Many Shades of Resistance

In what follows, I analytically describe the newsroom clash in RTVS which resulted in a significant staff turnover. Based on the interviews and the document analysis, I inductively developed a classification scheme (Marradi,

1990) for the resistance practices used by the journalists. They can be divided into two groups (Figure 1): resistance when a change within the organisation is deemed possible, and resistance when change within the organisation is no longer deemed possible. Another dividing line establishes resistance practices that can be either internal (i.e., intra-organisational) or external (i.e., extra-organisational). The dissatisfied journalists at first believed that change was possible and, as a general rule, although with some exceptions, they first used internal resistance practices, then they resorted to external-resistance practices. These practices did not lead to change, so the journalists eventually accepted the status quo and opted for resistance practices based on personal reactions to the unsatisfactory situation. Although Figure 1 suggests a general direction of individual steps, the process is reiterative rather than linear (e.g., requesting a meeting with the management was a step that was taken repeatedly in several stages of the conflict).

4.1. Prelude and Exposition: The Election of a New Director General

The prelude to the newsroom conflict at RTVS started with the election of the new director general. In a secret ballot in June 2017, the Slovak parliament chose Jaroslav Rezník to take the top position at RTVS. This choice was received with apprehension by part of the journalistic community and the public for two reasons. First, in the months preceding the election, the two leading politicians, then-Prime Minister Robert Fico (the chairman of Smer-SD party) and then-parliament speaker Andrej Danko (the chairman of the Slovak National Party), persistently and openly criticised RTVS and its management for being anti-Slovak, anti-government, biased, unfair with questions, and not publicising the successes of the government (e.g., Benedikovičová, 2016), and they claimed that "there must be a change in the leadership of RTVS" ("IPI criticises Slovak PM's," 2017). The second reason was personal: Jaroslav Rezník, the new director general, was believed to have ties to the Slovak National Party, the party that helped significantly to push through his nomination. Thus, the new director general took office in August 2017 in an atmosphere of tension and negative expectations.

The concerns were fuelled by the changes in the top management in September 2017. Rezník broke one of the key unwritten rules of journalism—to maintain a strict border between journalism and public relations—and appointed three former press officers

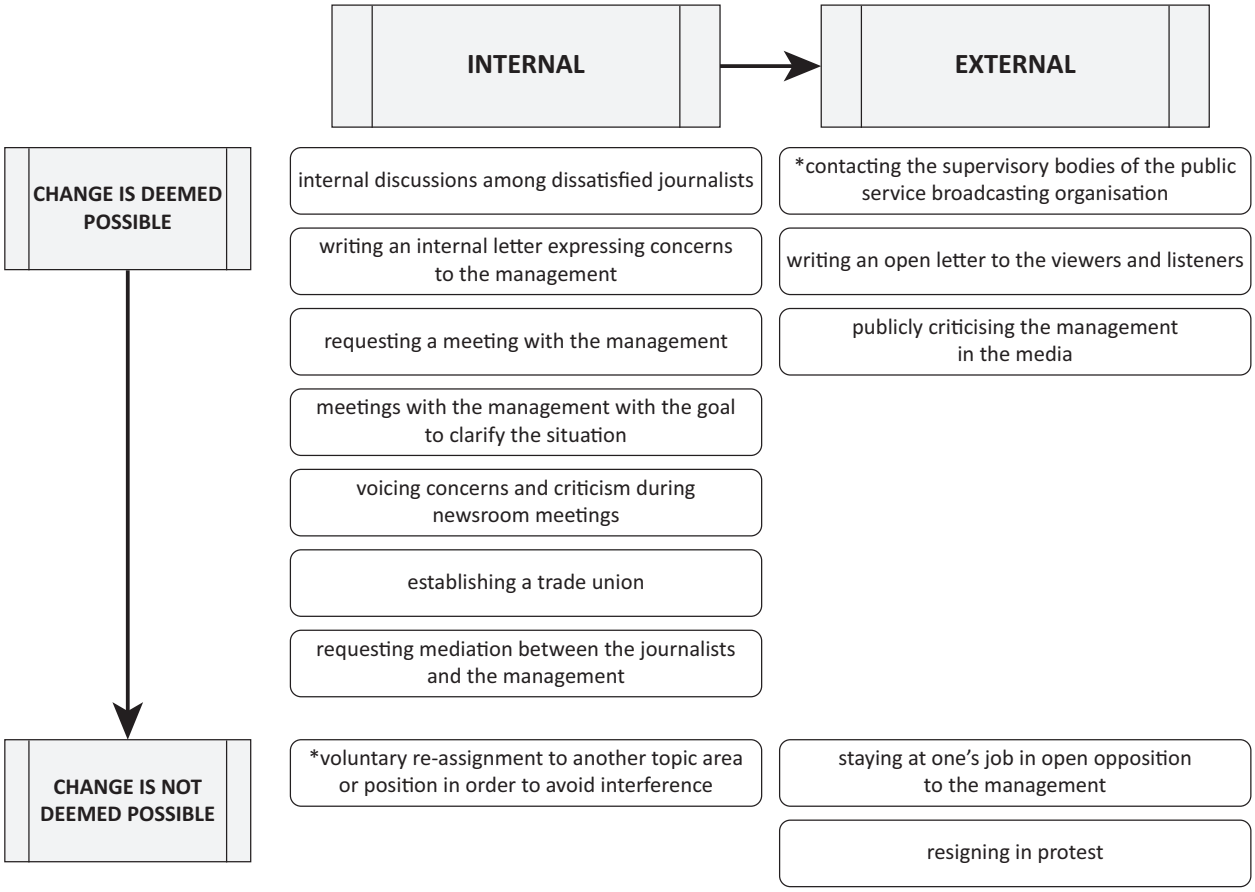

Figure 1. Resistance practices of the journalists facing internal interference with their professional autonomy. Note: The practices marked with an asterisk were mentioned by the research participants as a considered option, but they were not actually used.

from ministries and state organisations as the new top-managers for the television and radio newscasts. This was seen by some journalists as a conflict of interest that put RTVS' credibility in danger. The new director general refused to acknowledge the conflicts of interest and take steps to ensure that the people coming from the political environment would not have the direct control of the news content. As a consequence, the editor-in-chief resigned in protest, and the rest decided to wait and see what would come; at this stage, the concerned journalists discussed the issue internally and informally.

4.2. Collision: A Shutdown of an Investigative Programme and an Explosive Staff Meeting

In winter 2018, the management of RTVS decided to shut down its only investigative TV programme *The Reporters* (*Reportéri* in Slovak) without any discussion with the show's editors, ostensibly due to the lack of quality. This decision came shortly after the programme broadcast a story critical of Matica Slovenská, a state-funded national cultural organisation linked to the Slovak National Party and to Jaroslav Rezník, the director general, whose father happened to be a member of the organization's presidium. After protests from both the media community and the public, the show was reintroduced in September 2018, but the production team had been changed. Furthermore, soon after airing the Matica Slovenská story and the subsequent complaint of the organisation, RTVS broadcast what was marked by several interviewed journalists as an unusually laudatory story about the organization in its evening news programme.

The journalists tried to resolve their concerns within the organisation. They wrote an internal (non-public) letter to the management and requested a meeting for management to answer their questions:

> In January, the newsroom became very concerned about this, but we talked internally that we didn't want to be hysterical and that maybe these were just communication misunderstandings that we wanted to resolve within the organisation. We wrote an internal letter signed by 77 or 78 people. [It] was a short text with these points that were very worrying….[We wrote] that we believe that it was a misunderstanding and that we were asking Director General Rezník for an immediate staff meeting to explain all of this. (Interview with an RTVS reporter, July 2018)

The staff meeting took place in January 2018 and included approximately 70 reporters from both the television and radio news divisions and the management that was supposed to calm the situation. The most noteworthy moment, according to various research participants, was a speech by one of the radio managers. Allegedly, he explained, the laudatory news story about Matica Slovenská was broadcast because, as cited by several of the interviewed reporters, "sometimes steps need to be taken to soften the impact of a complaint" (Interview with an RTVS reporter, February 2019). This was met with loud protests from the reporters—they described their reaction as shock, some of them because they found such an approach to be journalistically and morally unacceptable and others because they were surprised that he admitted it so openly. As summarized by one of the interviewed reporters, "[E]veryone was shocked that he said that out loud. Even if it does happen, one does not say it [laughter]" (Interview with an RTVS reporter, February 2019).

Soon after the meeting, two of the three new top managers (i.e., former spokespersons) decided to resign, officially for personal and health reasons. According to the interviewees, they left due to the growing tensions between the director general and the news staff in which they were caught. The tension grew when one of the vacant managerial positions—the head of television and radio news and current affairs—was taken by Vahram Chuguryan, a journalist-turned-spokesman who worked, among other things, as a spokesperson for two ministers who were nominated by the then-coalition parties. Again, his professional past raised concerns among some reporters.

4.3. Crisis: Banned Badges, Quarrels About Objectivity, Open Letters, and Layoffs

The conflict entered its most heated stage in the strained atmosphere that followed the murder of Ján Kuciak, an investigative journalist who was shot dead with his fiancée in their home at the end of February 2018. Vahram Chuguryan, the new top-manager, asked the television reporters not to wear badges with the portrait of the murdered couple on-screen. The badge was originally designed and produced by the publisher of the news website for which Ján Kuciak had worked, and it quickly became a symbol of solidarity with the victims and their families. The badge became popular with many of the participants of the anti-government mass protests that were sparked by the murder and that eventually led to the resignation of Prime Minister Robert Fico and his cabinet. The new top-manager considered the badge to be a political symbol. Shortly after a discussion in the newsroom, the journalists leaked the story about the "ban on badges" to their colleagues in other media. At this stage, the dissatisfied journalists started to publicly express their concerns and criticise the management.

The tension grew into an open conflict. In the interviews, both sides mentioned hostile and stressful morning editorial meetings with every day quarrels, fights, shouting, and insults. It seems that what followed was a clash of two groups with two distinct journalistic cultures and different political orientations (including foreign political orientation in terms of East vs. West). On a practical level, this led to arguments about when a balancing quote was needed, what to include as the

"opposite'' view, what sources to quote, who was a credible and relevant expert source, and how to name certain events. In short, these questions concerned objectivity, and even more broadly, they concerned the perceptions of journalistic roles and of public service. Also important to mention is that, according to some of the interviewed journalists, these clashes often had one thing in common—they were somehow related to the Slovak National Party (i.e., the party that pushed through the nomination of the director general), its leaders, and the areas of its interest.

None of the interviewed reporters mentioned any case of censorship. Nobody reported any direct command or prohibition with regard to the production of journalistic content, nor any case of an editorial change without the author´s consent. However, all the reporters, regardless of whether they decided to leave RTVS, complained of pressure exerted on them or their colleagues to conform to the views of the managers and to stop challenging them. According to the interviewed reporters, punishment was meted out for failures to adjust to their managers' perception of objectivity, their notion of how to select sources, and the actors and opinions to include through two disciplinary measures: excessive negative feedback and cuts to bonuses, which were otherwise a significant part of their monthly pay.

A step that some of the journalists considered, but decided not to take, was contacting the supervisory body of RTVS—the RTVS Council. The Council oversees the functioning of RTVS, sets the salary of the director general, and it can submit a motion for filing a proposal for their dismissal to the relevant committee of the parliament. All nine council members are elected directly by the parliament. However, the interviewed journalists did not turn to the Council because they did not believe that the members were genuinely interested in discovering what was going on in the newsroom (and the members of the Council did not approach them on their own).

The journalists continued to request a meeting with the director general to describe the situation in the newsroom. However, according to the interviewed reporters, despite an initial promise made after the first big staff meeting in January, he avoided the meetings, postponed them, or cancelled one at the last minute. In this situation, in April 2018, around 60 RTVS reporters and editors signed a critical open letter to the viewers and listeners. They stated that they "continue to work freely… but in a hostile climate" and that they fight with "distrust in [their] superiors, their intentions, and their skills" ("Otvorený list," 2018). They objected to the conflicts of interest of their managers and former spokespersons, and they accused the management of suppressing critical voices. In their public response, Jaroslav Rezník and the top managers labelled the signatories as young, inexperienced, and radical. Also, in reaction to the open letter, 35 reporters, anchor-persons, and other RTVS employees signed another open letter stating that they work freely and without pressure.

Soon after, the RTVS managers ended the contracts of four reporters who were among the signatories of the critical open letter and who openly confronted and criticised their superiors during the editorial meetings. The management took advantage of the fact that, due to the lack of financial resources, several of the RTVS personnel were technically self-employed contractors, even though they were a stable and long-term part of the newsroom. Therefore, when getting rid of undesirable personnel, it was possible to cancel their contracts overnight instead of providing a much more complicated formal dismissal notice.

4.4. Peripety: The Establishment of a Trade Union and Requests for Mediation

In April 2018, the opposing journalists established a trade union to protect their jobs and compel management to hold a meeting. According to Slovak law, the employer is obliged to negotiate with a trade union. Moreover, it is illegal to dismiss employees' representatives during their term of office and for six months afterwards. Also, the employer may give notice or immediately terminate the employment of a member of a trade union only with the prior consent of the employee's representatives.

At the same time, there were repeated attempts at independent mediation to resolve the conflict. These attempts failed as well. Both sides of the dispute were not able to agree on the mediator: The management wanted to select the mediator and expected the journalists to accept the choice.

4.5. Catastrophe: A Wave of Resignations

Taking into account the aforementioned events, it came as no surprise that at the end of May 2018, 12 television reporters and editors handed in their notices of termination. Resignation in protest happened shortly after the unsuccessful negotiation between the trade union with the director general. According to one of the reporters, before the meeting they still hoped that he did not have enough information about what was going on in the newsroom: "After this meeting we understood that he knew what was going on, he knew that our letters of resignations were almost on the table, but he actually did not care" (Interview with an RTVS reporter, July 2018). Thus, the non-extension of the contracts for the four journalists, together with a (perceived) lack of interest on the part of the general director, seemed to be the turning point after which journalists stopped trying to change the conditions within the organisation and started to focus on resolving their personal situation within the status quo.

When explaining their decision, none of the reporters mentioned any experience of censorship but several of them expressed that the pressure from the new management put them at risk of self-censorship:

> The situation was terribly bad and there was enormous pressure for self-censorship. That's the main thing. No one told us what to write, but everyone, subconsciously, was already considering that I know that if I write this there, I know that they [the managers] will criticise me harshly tomorrow morning. (Interview with an RTVS reporter, February 2019)

As summarized by one of the interviewed reporters: "Censorship does not manifest itself only by someone saying that you can't broadcast something, but also by the exemplary punishing of colleagues who get bogged down in a topic that the management is not comfortable with" (Interview with an RTVS reporter, October 2018). As the reporters and editors who decided to resign explained in their public statement, they repeatedly tried to come to an agreement with the management but had now reached a point beyond which they could not go further, and that their resignation was a matter of professionalism and journalistic honour. They also declared: "We loved, love, and will love RTVS….We have children, mortgages, and obligations, but most of all—we love this job very much. Under these circumstances at RTVS, however, we cannot carry out [our work] as faithfully as before" (Spravodajské odbory RTVS, 2018). The management of RTVS stated that they regretted this decision but respected the choice of the reporters.

Besides resignation in protest, the dissatisfied journalists had two other options. One of the research participants considered voluntary re-assignment to another topic area or position in order to avoid the pressure (e.g., a transfer from the news section to the current affairs section), but in the end, she decided otherwise. Another option was to stay at one's job, but in open opposition to the management. This included open criticism both within the organisation and publicly. Some of the interviewed journalists who at some point belonged to the rebel group opted for yet another approach: They decided to focus on their own jobs and retreated into "internal exile"; however, this could be considered to be a strategy of acceptance rather than a strategy for resistance.

By Summer 2019, around two-thirds out of 26 television reporters left or were denied contract extensions. The same applies to roughly one-half of the editors and all of the on-line editors.

5. Conclusions

This article explored the resistance practices used by the journalists who worked for RTVS, the Slovak public service broadcaster, to cope with what they perceived as the undue and journalistically unwarranted interference of their superiors with their professional autonomy. The resistance practices identified in this case study included internal discussions among dissatisfied journalists, the writing of an internal letter to express concerns to the management, requests for a meeting with the management, meetings with the management with the goal to clarify the situation, the voicing of concerns and criticism during regular newsroom meetings, contacting supervisory bodies of the public service broadcasting organisation, the establishment of a trade union, requests for independent mediation between the journalists and the management, the writing of a public open letter to the viewers and listeners, the public expression of concerns and criticism in the media, requests for voluntary re-assignment to another topic area or position in order to avoid pressure, staying at one's job but in open opposition to the management, and resignation in protest. To summarize, the journalists gradually moved from voice to exit, indicating that—in line with past studies (Borden, 2000; Davidson & Meyers, 2016)—using voice eventually comes at extremely high costs in journalism.

These resistance practices are very different from those identified in previous research, which largely examined cases of external political interference (Ataman & Çoban, 2019; Barrios & Miller, 2020; Slavtcheva-Petkova, 2019) or external/internal commercial interference (Borden, 2000; Goyanes & Rodríguez-Castro, 2019; Hanusch et al., 2017). This suggests that the choice of specific types of resistance is closely related to the source and type of interference. Another important variable is the type of organisation; some of the identified resistance practices can only be used in public service media (e.g., contacting supervisory bodies of public service broadcasting organisation).

In the organisational studies literature, voice strategies are distinguished from the employer's perspective based on active-passive and constructive-destructive dimensions (Gorden, 1988). However, this study argues that professional conflicts should not be assessed solely from the standpoint of the employer. If, for instance, voice has to be used to defend professional autonomy, from the viewpoint of the journalistic profession, labelling individual strategies as constructive or destructive may look quite different. This study therefore introduces a different conceptualisation that adopts the perspective of employees, and that is based on two main dimensions, reach and aim. In terms of reach, resistance practices can be either internal (i.e., used within the organisation) or external (i.e., extending beyond the boundaries of the organisation). Also, they can aim at the improvement of the conditions within the organisation (i.e., these practices are used when journalists still believe change is possible) or they can aim at resolving the personal situation when change in the organisation is no longer considered possible. In the first stage, the dissatisfied journalists tried to improve the conditions in the organisation with internal resistance practices. As a general rule, although with some exceptions, they only resorted to external resistance practices once they had exhausted the internal ones. In the second stage, the journalists accepted the status quo and opted for resistance practices that were focused on personal reactions to the unsatisfactory status quo. The reiterative model

of resistance practices, including the general direction of individual steps, introduced in this study can be further developed and tested in other newsroom conflicts both in public service and commercial media organisations.

Also, this case study supports Hirschman's (1970) claim that the key moderating variable that influences the response to dissatisfaction is loyalty. The journalists had long sought to change the situation in the organisation because they identified with its values and mission, felt to be a part of it, and considered the steps of the new management to be a threat to its reputation and ability to fulfil its mission. Another important variable that was identified in this case study is peer support and the journalists' ability to organise and resist collectively. Several of the resistance practices were collective (not individual), and many of the interviewed journalists mentioned that mutual support was a significant factor that helped them continue in resistance. The significance of the collective aspect of resistance must be emphasized all the more as journalists are often reluctant to organise themselves and practice their occupational voice (Davidson & Meyers, 2016).

Finally, this case study suggests that, even though journalistic autonomy historically did not belong to the core ideal-typical values of journalism in Slovakia, a country that once belonged to the Soviet sphere of influence, a significant part of the journalists who worked at RTVS considered it to be an essential value that was worth defending, even at the cost of their jobs. This is an important finding as throughout its history, public service broadcasting in Slovakia has been repeatedly attacked and used as a political tool (most flagrantly under Prime Minister Vladimír Mečiar and his autocratic style of government in 1994–1998). In the previous cases, journalists either adapted to the new conditions or decided to leave without much struggle. This time, the dissenting journalists refused to succumb to self-censorship or to adopt the practice of "adekvatnost" (i.e., a state of being adequate) that produces journalism corresponding to the authorities' expectations (Schimpfossl & Yablokov, 2014), which is known from contemporary Russia. Thus, even if some scholars argue that, in countries where the development of journalism culture has been disrupted, including Slovakia, journalistic autonomy is less deeply rooted in professional ideology or not regarded as important at all (Lauk & Harro-Loit, 2016), the conflict at RTVS signals that a non-negligible part of Slovak journalists abandoned the Soviet model of journalism and adopted the key values of the Western journalism culture not only rhetorically but also through their actions. This is a positive signal in terms of the vitality of the journalistic profession and the strength of public service media in Slovakia and, more broadly, in the CEE region.

Acknowledgments

The author wishes to thank all the research participants for their time and insight, as well as to the issue editors and two anonymous reviewers for helpful comments on earlier draft of the manuscript.

Conflict of Interests

The author declares no conflict of interests.

References

Ahva, L., van Dalen, A., Hovden, J. F., Kolbeins, G. H., Löfgren Nilsson, M., Skovsgaard, M., & Väliverronen, J. (2017). A welfare state of mind? Nordic journalists' conception of their role and autonomy in international context. *Journalism Studies*, *18*(5), 595–613.

Akhrarkhodjaeva, N. (2017a). Russian media and journalists' dilemma between "exit, voice, and loyalty." *Russian Analytical Digest*, *197*, 5–8.

Akhrarkhodjaeva, N. (2017b). *The instrumentalisation of mass media in electoral authoritarian regimes: Evidence from Russia's presidential election campaigns of 2000 and 2008*. ibidem-Verlag.

Aravopoulou, E., Mitsakis, F. V., & Malone, C. H. (2017). A critical review of the exit-voice-loyalty-neglect literature: Limitations, key challenges, and directions for future research. *The International Journal of Management*, *6*(3), 1–10.

Ataman, B., & Çoban, B. (2019). Turkey: How to deal with threats to journalism? In E. Eide, K. Skare Orgeret, & N. Mutluer (Eds.), *Transnational othering—Global diversities: Media, extremism, and free expression* (pp. 171–190). Nordicom.

Barrios, M. M., & Miller, T. (2020). Voices of resilience: Colombian journalists and self-censorship in the post-conflict period. *Journalism Practice*. Advance online publication. https://doi.org/10.1080/17512786.2020.1778506

Benedikovičová, M. (2016, November 26). Sme v krčme? Dočerta, spamätajte sa, útočil Fico na moderátora, keď sa pýtal na Lajčáka [Are we in a pub? Damn, get real, Fico attacked the moderator when he asked about Lajčák]. *Denník N*. https://dennikn.sk/617975/sme-v-krcme-docerta-spamatajte-sa-utocil-fico-na-moderatora-rozhlasu-ked-sa-pytal-na-lajcaka

Borden, S. L. (2000). A model for evaluating journalist resistance to business constraints. *Journal of Mass Media Ethics*, *15*(3), 149–166.

Bourdieu, P., & Wacquant, L. J. D. (1992). *An invitation to reflexive sociology*. Polity Press.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101.

Clark, M., & Grech, A. (2017). *Journalists under pressure: Unwarranted interference, fear, and self-censorship in Europe* (1st ed.). Council of Europe.

Davidson, R., & Meyers, O. (2016). "Should I stay or should I go?": Exit, voice, and loyalty among journalists. *Journalism Studies*, *17*(5), 590–607.

Deuze, M. (2005). What is journalism? Professional iden-

tity and ideology of journalists reconsidered. *Journalism*, *6*(4), 442–464.

Dragomir, M. (2017, August 29). *The state of Hungarian media: Endgame*. LSE Media Policy Project. https://blogs.lse.ac.uk/medialse/2017/08/29/the-state-of-hungarian-media-endgame

Dzięciołowski, K. (2017). *Is there a chance for non-partisan media in Poland?* (Reuters Institute Fellowship Paper). Reuters Institute; University of Oxford. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2017-12/Is%20there%20a%20chance%20for%20non-partisan%20media%20in%20Poland%20-%20Krzysztof%20Dzieciolowsk%20Paper.pdf

Farrell, D. (1983). Exit, voice, loyalty, and neglect as responses to job dissatisfaction: A multidimensional scaling study. *Academy of Management Journal*, *26*(4), 596–607.

Fedirko, T. (2020). Self-censorships in Ukraine: Distinguishing between the silences of television journalism. *European Journal of Communication*, *35*(1), 12–28.

Freidson, E. (1994). *Professionalism reborn: Theory, prophecy, and policy*. University of Chicago Press.

Gorden, W. I. (1988). Range of employee voice. *Employee Responsibilities and Rights Journal*, *1*(4), 283–299.

Goyanes, M., & Rodríguez-Castro, M. (2019). Commercial pressures in Spanish newsrooms: Between love, struggle, and resistance. *Journalism Studies*, *20*(8), 1088–1109.

Hanretty, C. (2011). *Public broadcasting and political interference*. Routledge.

Hanusch, F., Hanitzsch, T., & Lauerer, C. (2017). "How much love are you going to give this brand?" Lifestyle journalists on commercial influences in their work. *Journalism*, *18*(2), 141–158.

Hiltunen, I. (2019). Experiences of external interference among Finnish journalists. *Nordicom Review*, *40*(1), 3–21.

Hirschman, A. O. (1970). *Exit, voice, and loyalty: Responses to decline in firms, organizations, and states*. Harvard University Press.

Hughes, S., Garcés, M., Márquez-Ramírez, M., & Arroyave, J. (2017). Rethinking professional autonomy: Autonomy to develop and to publish news in Mexico and Colombia. *Journalism*, *18*(8), 956–976.

IPI criticises Slovak PM's remarks on public broadcaster. (2017, April 12). International Press Institute. https://ipi.media/ipi-criticises-slovak-pms-remarks-on-public-broadcaster

Jančáriková, T. (2018, May 31). Reporters quit Slovak public broadcaster to protest at political pressure. *Reuters*. https://www.reuters.com/article/us-slovakia-politics-media-idUSKCN1IW25M

Kassing, J. W. (2002). Speaking up: Identifying employees' upward dissent strategies. *Management Communication Quarterly*, *16*(2), 187–209.

Koivunen, A. (2017). #Sipilägate and the break-up of the political bromance: Crisis in the relationship between Finnish media and politicians. *Nordicom*, *39*(1), 44–51.

Lauk, E., & Harro-Loit, H. (2016). Journalistic autonomy as a professional value and element of journalism culture: The European perspective. *International Journal of Communication*, *11*, 1956–1974.

Marradi, A. (1990). Classification, typology, taxonomy. *Quality and Quantity*, *24*(2), 129–157.

Nissen, C. S. (2014). Organisational culture and structures in public media management: In search of a model for the digital era? In M. Głowacki & L. Jackson (Eds.), *Public media management for the twenty-first century: Creativity, innovation, and interaction* (pp. 81–102). Routledge.

Nissen, C. S. (2016). Obeying his masters' voices: Managing independence and accountability in public service media between civil society and state. In G. F. Lowe & C. Brown (Eds.), *Managing media firms and industries: What's so special about media management?* (pp. 121–142). Springer.

Otvorený list členov Sekcie spravodajstva a publicistiky RTVS (plné znenie) [Open letter of the members of the news and current affair section to the viewers and listeners (full text)]. (2018, April 4). *SME*. https://domov.sme.sk/c/20795636/otvoreny-list-clenov-sekcie-spravodajstva-a-publicistiky-rtvs-plne-znenie-reznik.html

Reich, Z., & Hanitzsch, T. (2013). Determinants of journalists' professional autonomy: Individual and national level factors matter more than organizational ones. *Mass Communication and Society*, *16*(1), 133–156.

Reporters Without Borders. (2020). *Slovakia*. https://rsf.org/en/slovakia

Saldaña, M., Sylvie, G., & McGregor, S. (2016). Journalism–business tension in Swedish newsroom decision making. *Journal of Media Ethics*, *31*(2), 100–115.

Schimpfossl, E., & Yablokov, I. (2014). Coercion or conformism? Censorship and self-censorship among Russian media personalities and reporters in the 2010s. *Demokratizatsiya: The Journal of Post-Soviet Democratization*, *22*(2), 295–311.

Slavtcheva-Petkova, V. (2019). Fighting Putin and the Kremlin's grip in neo-authoritarian Russia: The experience of liberal journalists. *Journalism*, *20*(11), 1530–1546.

Spravodajské odbory RTVS [News service trade union RTVS]. (2018, May 31). *Dnes, 31. mája 2018, sme sa dvanásti rozhodli podať výpovede* [Today, May 31, 2018, twelve of us have decided to resign] [Facebook status update]. https://www.facebook.com/watch/?v=1870199999707113&ref=sharing

Standaert, O., Hanitzsch, T., & Dedonder, J. (2019). In their own words: A normative-empirical approach to journalistic roles around the world. *Journalism*. Advance online publication. https://doi.org/10.1177/1464884919853183

Subhakaran, S. E., & Dyaram, L. (2018). Interpersonal

antecedents to employee upward voice: Mediating role of psychological safety. *International Journal of Productivity and Performance Management*, *67*(9), 1510–1525.

The Economist Intelligence Unit. (2020). *Democracy index 2019: A year of democratic setbacks and popular protest*. https://www.in.gr/wp-content/uploads/2020/01/Democracy-Index-2019.pdf

Tunstall, J. (1971). *Journalists at work. Specialist correspondents: Their news organizations, news sources, and competitor-colleagues*. Constable.

Weaver, D. H., Beam, R. A., Brownlee, B. J., Voakes, P. S., & Wilhoit, G. C. (2007). *The American journalist in the 21st century: US news people at the dawn of a new millennium*. Routledge.

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE.

About the Author

Marína Urbániková is an assistant professor at the Department of Media Studies and Journalism, Masaryk University, Czech Republic. Her research interests include topics of journalistic autonomy, journalism cultures, public service media and its independence, security and safety of journalists, journalism education, and public confidence in media and journalists. She has published, among others, in *European Journal of Communication*, *Journalism Practice*, and *International Communication Gazette*.

COGITATIO

Media and Communication (ISSN: 2183–2439)
2021, Volume 9, Issue 4, Pages 104–115
https://doi.org/10.17645/mac.v9i4.4217

Article

Protest Event Analysis Under Conditions of Limited Press Freedom: Comparing Data Sources

Jan Matti Dollbaum [1,2]

[1] SOCIUM Research Center on Inequality and Social Policy, University of Bremen, Germany;
E-Mail: dollbaum@uni-bremen.de
[2] Research Center for East European Studies at the University of Bremen, Germany

Submitted: 25 February 2021 | Accepted: 28 May 2021 | Published: 21 October 2021

Abstract
The investigation of long-term trends in contentious politics relies heavily on protest event analysis based on newspaper reports. This tends to be problematic in restricted media environments. To mitigate the effects of bias and (self-)censorship, researchers of protest in authoritarian regimes have experimented with other sources such as international media and dissident websites. However, even though classical news media are easier targets for repression, journalistic reports might still outperform other sources regarding the quality of information provided. Although these advantages and disadvantages are known in the literature, different types of sources have seldom been tested against each other in an authoritarian context. Using the example of Russia between 2007 and 2012, the present article systematically compares protest event data from English-language news agencies, dissident websites, and several local sources, first and foremost with a view to improving methodological knowledge. The analysis addresses broad trends across time and space as well as the coverage of specific regions and single protest events. It finds that although the data sources paint different pictures of protest in Russia, this divergence is systematic and can be put to productive use. The article closes with a discussion on how its findings can be applied in other contexts.

Keywords
authoritarian regimes; media freedom; opposition; protest event analysis; Russia

Issue
This article is part of the issue "Media Control Revisited: Challenges, Bottom-Up Resistance and Agency in the Digital Age" edited by Olga Dovbysh (University of Helsinki, Finland) and Esther Somfalvy (Research Centre for East European Studies at the University of Bremen, Germany).

1. Introduction

Protest event analysis is one of the most important and widespread methods to investigate social movements and protest cycles on a large scale; however, the insights drawn from it can only be as good as the data on which it is based. Traditionally, protest event analysis relies on journalistic reporting, which is usually reliable (if trustworthy sources are selected), but it nevertheless comes with certain biases concerning the selection of events and the reported details (Earl et al., 2004; Gladun, 2020).

In political environments where journalists cannot operate freely, the usefulness of protest event analysis data from news sources is questionable. At least, researchers must take additional precautions to circumvent or mitigate the additional biases that are introduced through (self-)censorship. Some protest data sets, therefore, rely only on international news agencies (Weidmann & Rød, 2019). This approach, however, has the disadvantage that it captures only the internationally visible fraction of the protest landscape, which may not be representative of protest as a whole. Other approaches include: (1) selecting as many different local sources as possible to minimize bias of individual sources—the so-called "blanketing strategy" (Beissinger, 2002); (2) using social media data (Zhang & Pan, 2019a); and (3) relying on dissident websites (Lankina, 2015; Robertson, 2013). While all of them have their own

advantages, they also have drawbacks: The blanketing strategy is highly resource intensive, social media posts often do not contain easily accessible information on important details such as topic and size (Zhang & Pan, 2019a), and activist data can be politically biased.

In this article, I undertake one of the first attempts to systematically compare different sources of protest event data, addressing broad trends over time and space, their overlap, and their thematic coverage. On the example of protest in Russia between 2007 and 2012, I compare data sourced from international news agencies (Weidmann & Rød, 2019), data extracted from dissident websites (Lankina, 2018; Robertson, 2013), and data from various local sources, including journalism and official accounts (Semenov, 2017). This is not an effort to expose flaws and biases in the different data sets. Instead, it is an attempt to better understand whether such biases systematically relate to the sources and structure of data sets, and to examine which data source is best for answering which type of research question.

Russia lends itself well to such an undertaking: First, in the period studied, it represented a paradigmatic case of a modern hybrid regime that combined democratic elements such as multi-party elections with obstruction, targeted repression, and manipulation—skewing the institutional playing field toward the political leadership (Levitsky & Way, 2010). This hybridity included the media sphere: While there was no official censorship, there were highly visible cases of political pressure through ownership changes and repression of journalists (McFaul & Stoner-Weiss, 2008). Reporters thus needed (and still need) to navigate a complex set of intentionally vague, unwritten rules which may result in self-censorship on sensitive topics such as protest. Moreover, there was considerable subnational variation of repression against activists and media (Dollbaum, 2020a; Petrov & Titkov, 2013). The Russian case can thus also be helpful to derive more general hypotheses on the interplay of press freedom, political activism, and the quality of protest event data.

Overall, the analysis finds that the data sources differ from each other in the picture they paint of protest in Russia. This divergence is, however, systematic, meaning that it can be put to productive use by matching one's data with the research question: While international media are best in capturing large protest waves, activist-based data quite consistently document regional trends. Finally, for case study research that focuses on single events, there seems to be no way around utilizing a diverse set of local sources.

2. The Evolution of Protest Event Analysis

Protest event analysis as a method evolved over several decades. Hutter (2014) identified four "generations" in this evolution: The first comprises early works by, for instance, Tilly et al. (1975) that established the method of using reports on protest events (usually newspapers) to map popular contention across time and space. Fueled by nascent debates on source selection, a second generation (Kriesi et al., 1995) paid greater attention to sources and, in addition, increasingly used the technique for cross-national comparisons. A third generation then more systematically addressed selection bias within newspaper sources themselves. Predictors for newspaper coverage include the number of protestors, the degree of disruption and violence, the presence of counter-demonstrations, and police involvement (Earl et al., 2004). Others have found that other additional factors such as the timing in a legislative process (Oliver & Maney, 2000), the sponsorship of a protest by established social movement organizations (McCarthy et al., 2008), and media attention to similar events (Hellmeier et al., 2018) also make coverage more likely. Similar distortions apply to television coverage (Wouters, 2013). A fourth phase then went beyond single protest events or aggregated protest cycles (Tarrow, 1993) as the unit of analysis, instead analyzing "contentious episodes" (Kriesi et al., 2019). Stages three and four also saw the increased use of machine learning to classify relevant reports (Bremer et al., 2020) or to automatically annotate whole articles (Hanna, 2017), a technique that has the potential to greatly reduce the resources necessary to conduct protest event analysis.

3. Protest Event Analysis in Unfree Media Environments

Problems concerning the sampling and the quality of information in the source material are exacerbated when the context is characterized by limited freedom of the press. The reported studies refer to democratic contexts where bias usually results from sources' or journalists' political orientation or from competition for attention. Repressive contexts add censorship and self-censorship as another layer of bias (McCarthy et al., 2008). A partial solution lies in what Beissinger terms "a 'blanketing strategy,' utilizing multiple sources and multiple types of information whenever they are available" (Beissinger, 2002, p. 476).

This idea also underlies two approaches that expand the range of source types used for protest event analysis. In the strictly controlled media environment of closed authoritarian regimes such as China, a viable alternative to news reporting may be to gather data from social media. In a pioneering study, Zhang and Pan (2019a) found that social media data does indeed help to increase the share of otherwise underreported rural protest, but so far it has been possible to automatically identify only a very small set of variables, excluding important information such as "size, action form, claims/issues, targets, organizers, and violence" (Zhang & Pan, 2019b, pp. 76–77).

The second approach involves activist projects that document protest. These often work through a network of activist-correspondents who monitor their local

environment while a website aggregates that information. In Russia, two such projects have been turned into data sets: Lankina Russian Protest Event Dataset (LAruPED) and Institute of Collective Action (IKD). In the context of restricted media freedom, these activist-based data sets have the potential to be simultaneously more efficient and less affected by self-censorship than media-based data. However, given that they come from politically interested sources, they may also oversample specific topics at the expense of others. The following section sets up a research design that systematically compares four data sets of different origins.

4. Data Sources and Research Design

The empirical analysis will be conducted in two stages. In the first stage, I compare the Mass Mobilization in Autocracies Database (MMAD) data set that is based on international news reporting to two data sets derived from activist websites. All three cover the whole of Russia in varying time periods, with an overlap from March 2007 through to March 2012. The second stage then compares all three data sets with the Contentious Politics in Russia (CPR) data set gathered by Andrei Semenov in two Russian regions. The MMAD data are available on their own website, all other data sets are being uploaded to the Discuss Data project.

The analysis is conducted to show the different pictures of a country's protest landscape that different data sets produce, in order to better understand possible distortions when relying on a single data set of known origin (e.g., activist-based or international media-based protest data). As the goal is not to arrive at a substantive insight on protest dynamics in Russia but, instead, to fully reveal these cross-source differences, in the analysis I consciously do not make the data sets more comparable to each other (for instance, by subsetting) but let the differences in sources and construction produce their full effects.

I now briefly describe each data set regarding the origin, coverage, and coding criteria, and summarize the differences in Table 1. Based on these characteristics, I then derive hypotheses on how their respective pictures of protest in Russia differ from each other.

4.1. Mass Mobilization in Autocracies Database

The MMAD has been developed by a team around Nils Weidman at the University of Konstanz. It covers all authoritarian regimes, identified in accordance with Geddes et al.'s research (2014). The data is based on three international news agencies: Associated Press, Agence France Press, and BBC Monitoring. While these are all English-language sources, BBC Monitoring also translates local news, considerably enlarging the overall coverage (Weidmann & Rød, 2019, pp. 42–44). Nonetheless, the authors make clear that the data set may introduce bias, "as these agencies typically cater to audiences in Western countries and primarily report on events that are of some interest to them" (Weidmann & Rød, 2019, pp. 41–42). MMAD includes only protest events with an identifiable political motive. The definition, however, is broad, including all "matters of, or relating to, the government or the public affairs of a country" (Keremoglu et al., 2020, p. 2). Events include, for example, protests against the monetization of social benefits or a spike in fuel prices—claims that would be labelled social or economic by other authors (see Lankina & Voznaya, 2015). A far greater restriction, which is likely to affect the composition of the database, is introduced by MMAD's minimum threshold of 25 participants for an event to be included. MMAD also covers pro-government rallies, which, however, are excluded here.

4.2. Lankina/Lankina Russian Protest Event Dataset

This data set is the first of two that is based on activist websites. To compile it, a team around Tomila Lankina coded all entries on protest that were posted on the website namarsh.ru between 2007 and 2016 (Lankina & Tertytchnaya, 2020). The site is funded by the opposition politician Garry Kasparov, who led the cross-ideological opposition group "Other Russia." Its content comes from a Russia-wide network of correspondents. As Lankina writes, "the website is run by a team of activists sympathetic to the cause of the opposition and therefore interested in ensuring wide reporting and the most comprehensive harvesting of published online reports on protest" (Lankina, 2015, p. 30).

In contrast to MMAD, this data set excludes pro-government events (Lankina & Tertytchnaya, 2020, p. 22). Also, it does not have a lower limit on the number of participants. However, the authors are clear that the source does not represent full coverage of all protests (Lankina, 2015; Lankina & Voznaya, 2015). For instance, as the data set comes from an opposition website, "it may well contain a liberal bias in favour of pro-democracy activism" (Lankina & Tertytchnaya, 2020, p. 24). However, as of yet, there are no systematic comparisons with other data sources that could indicate where any such biases lie.

4.3. Reuter and Robertson/Institute of Collective Action

The third data set was constructed by Reuter and Robertson (2015; see also Robertson, 2013), covering the period from January 2007 through to March 2012. Its origin is a website, the IKD, in which a group of sociologists with sympathies for left-wing oppositional causes compile weekly reports on protest events across the country. As with LAruPED, IKD has no lower limit on the number of participants (but it does not record this number). Hence, the two activist-based data sets are similar structurally, but their political outlooks are quite different. While Kasparov's "Other Russia" blends liberalism with nationalist elements introduced by one of its constituent groups (the National Bolsheviks), the

group behind IKD "seeks to unite different social groups it describes as being 'without a voice,' " including "leftist groups, labour unions, and environmental and youth organizations" (Reuter & Robertson, 2015, p. 241).

As evident from these short descriptions, the two activist-based data sets come from very different sources. LAruPEd can be expected to oversample political protest, while the IKD data are likely to prominently cover social protest (when the two are compared directly, this expectation is confirmed, see Supplementary File). It is therefore not to be expected that they would produce exactly the same picture of protest. Instead, it will be the task of the empirical analysis to estimate the degree to which they produce similar pictures of protest dynamics despite the potential differences in political outlooks of their author collectives.

4.4. Semenov/Contentious Politics in Russia

The final data set, CPR, was collected by Andrei Semenov and colleagues at the Center for Comparative History and Politics at Perm State University. Here, I use two publicly available portions of it, covering the regions of Perm and Tyumen. The former is based on an online search with the Integrum service that archives over 40,000 Russian online, press, radio, and TV sources (Semenov, 2017). The latter is based on a comprehensive search of two local online media as well as participant observation. Crucially, it also uses official data—not records on actual protest events collected by police (as in the case of Robertson's data from the later 1990s; see Robertson, 2013), but data on applications submitted for protest events (in Russia, as in many other countries, protests that exceed one person need to be registered with authorities, which provides an excellent potential source of data—even though in most cases these data are inaccessible). The CPR data have no demonstrator threshold. Both approaches use several types of sources, approximating Beissinger's (2002) "blanketing strategy" in different ways and should thus decrease bias that results from source selection and selective coverage.

Table 1 compares the characteristics of all four data sets. The Supplementary Files contain a discussion of how the (few) duplicates in the data sets were dealt with.

4.5. Hypotheses

In the first stage of the empirical analysis, MMAD, LAruPED, and IKD will be compared to each other regarding broad trends in the distribution of protest. Comparisons include the distribution over time, the share of protests events in the regions (as opposed to Moscow and St. Petersburg), as well the rank order of regions according to the number of protest events. The difference in sources, as well as MMAD's 25-person threshold, make it likely that, compared to the other two, these data paint a different picture of protest in Russia. First, media coverage, in general, follows "issue attention cycles" (Downs, 1972), which have a bearing on protest coverage (Oliver & Maney, 2000) and thus potentially affect fully media-based data sources like MMAD (see also Gladun, 2020). Moreover, Herkenrath and Knoll (2011) have found substantial differences in protest coverage when comparing international and national news media sources—differences that I expect to show in the analysis:

> H1) Since MMAD is based on international media and imposes a 25-person threshold, it will focus on larger and more visible events, which results in a different distribution of events when compared to LAruPED and IKD.

Moreover, data extracted from reporting in international news media likely overrepresent events in the capital "where most foreign journalists have their workplaces" (Wüest & Lorenzini, 2020, p. 49). Therefore:

Table 1. Comparison of data sets used in analysis.

Data set	Sources	Topic of events covered	Includes pro-government?	Minimum participant threshold	Number of events in covered period[1]	Number of regions with at least 1 event in covered period[2]
MMAD news reports	International	Political (broad definition)	Yes (but excluded for analysis)	25	1,152	76
LAruPED	Activist website	All	No	none	4,497	76
IKD	Activist website	All	No	none	5,593	81
CPR	Multiple local sources	All	No	none	458	—[3]

Notes: [1] The covered period ranges from 16 March 2007 (the start date of the LAruPED) through 5 March 2012 (the end date of the IKD data). [2] Before the annexation of Crimea in 2014, the Russian state counted 83 subnational subjects, including "oblasts," "krais," "autonomous republics," and others. For convenience, these are all summarized under the label "regions" here. [3] The CPR data is compared to LAruPED and IKD only on the regions of Perm and Tyumen, see Sections 6.1 and 6.2.

H2) MMAD will show a greater concentration of events in Moscow and St. Petersburg compared to LAruPED and IKD.

Similarity in source type, on the other hand, should increase convergence—even if the specifics of the sources vary:

H3) Because LAruPED and IKD both come from local activists they will, overall, show greater convergence with each other than either of them does with MMAD.

In the second stage, the activist-based data sets will be compared to the CPR data that likely come closer to full coverage of all protest as it approximates Beissinger's "blanketing strategy" (2002, p. 476) to different degrees, leading to the following hypotheses:

H4) The CPR data in Perm and Tyumen cover more protest events than even LAruPED or IKD.

H5) Since the Tyumen data include a greater variety of source types, when compared to LAruPED and IKD, they cover more unique events than do the CPR data on Perm when compared to LAruPED and IKD.

Finally, I code protesters' claims as reported in CPR to match LAruPED, assigning each event one or more of six broad categories: political, economic, social, legal, ecological, and cultural (Lankina & Voznaya, 2015, p. 332; see Supplementary Files for coding details). This offers the chance to compare the CPR data to LAruPED regarding the thematic coverage of protest. As Lankina and Tertytchnaya (2020) have pointed out, LAruPED may be biased in favour of political protest, leading to the final hypothesis:

H6) LAruPED oversamples political protest, which leads to a higher share of political protest compared with CPR.

5. Empirical Analysis I: General Trends

5.1. Development Over Time

First, I graphically explore the distribution of protest across time. As expected, Figure 1 shows that MMAD differs strongly from the other two, with protest peaking in late 2011 and early 2012. This was the time of the country-wide For Fair Elections (FFE) protests—the largest and geographically most widespread protests in post-Soviet history—which attracted comprehensive

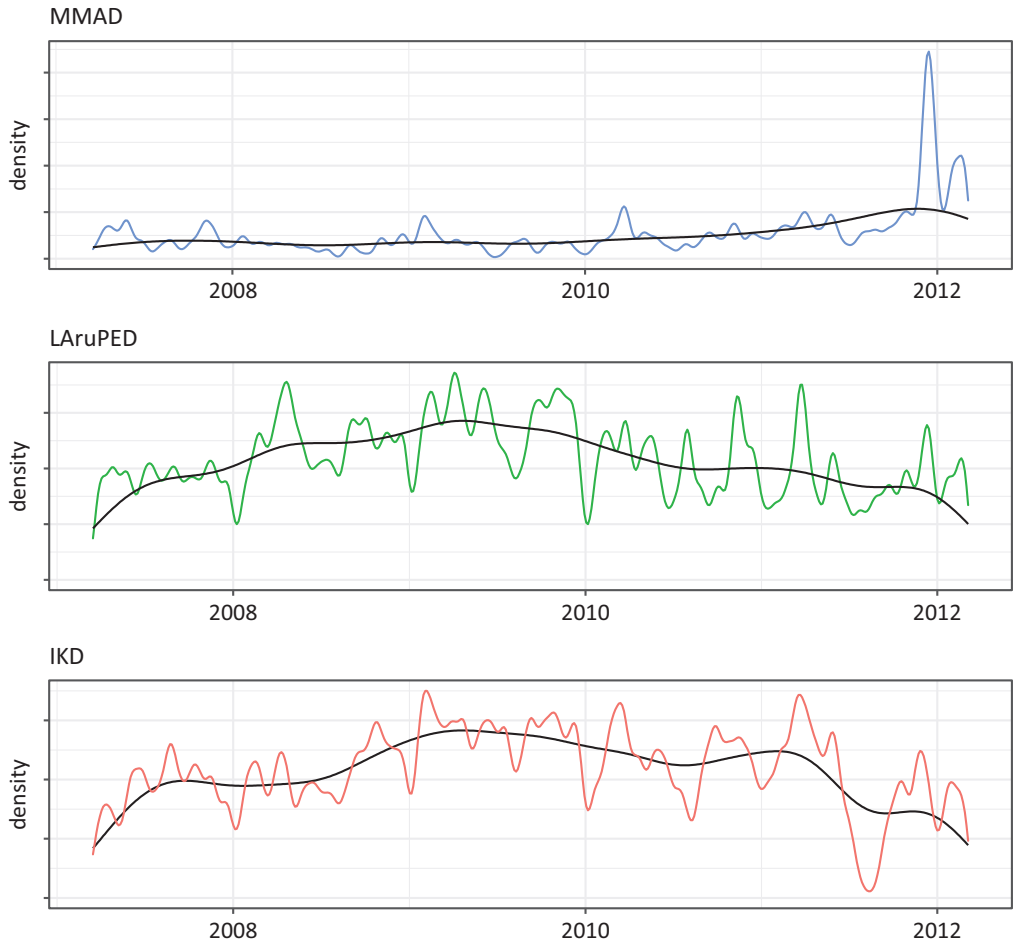

Figure 1. Protests events over time by data source, with density curve and trend line (March 2007–March 2012).

international media attention. But even when bracketing out this specific historical period, the distribution of MMAD diverges from that of LAruPED and IKD (see Figure A1 in the Supplementary Files). The latter two, by contrast, seem to depict broadly similar trends, confirming H1.

In the Supplementary Files, I show that the trend in the LAruPED data does not change markedly when the data are subsetted to only those events with 25 and more participants, which make them structurally more similar to the MMAD data (Figure A2). This suggests that differences between MMAD and the activist-based data sets do not just result from the different inclusion threshold.

5.2. Share of Regional Protest Events

The different trends observed above may, among other things, be because international media coverage underestimates the share of protest in the provinces (Gabowitsch, 2016). Figure 2 supports this assertion, demonstrating that MMAD features a substantially lower share of regional protest than both LAruPED and IKD. In 2011, however, the picture changes, with the share of regional protest in MMAD strongly increasing. This is likely once more connected to the FFE protests which, having begun in Moscow, quickly spread through the regions. This is supported when the period of the FFE protests is removed, which causes the share of regional protest in 2011 to drop substantially in all three data sets (see Figure A1 in the Supplementary File).

These observations suggest two things. First, the trend of concentration of protest in the capitals during the second half of the 2000s, as diagnosed by Robertson (2013), can be replicated with the LAruPED data (see also Lankina, 2015) and MMAD. This is a welcome finding as it points to the robustness of a broad trend that has become standard knowledge in protest research on Russia.

Second, there are two possible explanations for the sudden increase in the share of regional protest with the onset of the FFE protests. In one interpretation, the share surges because the FFE events are larger than previous regional protest so that more regional events cross MMAD's threshold of 25 participants. In the second explanation, the beginning of the protests wave triggered an "issue attention cycle," increasing the interest of international news agencies in regional protest. In other words, what used to be in the periphery of attention suddenly moved to the centre, which would mean that MMAD substantially underreported regional protest before the FFE protests. Both explanations are plausible. To test which has greater empirical support, I compute the average number of participants in regional protest events in LAruPED before and during the FFE period. If regional protest events did indeed become larger, this should be reflected in the LAruPED data as well.

The data show that regional protest events were indeed, on average, larger in the FFE period than they had been before: The median before December 4, 2011, is 100 participants. Between that date and the presidential elections on March 4, 2012, it doubles to 200. In part, this appears to be because more events cross the 25-participant threshold: Before the FFE protests, 24% of events are below that number; during FFE, that share drops to 9% (see table A1 in the Supplementary Files). Although these numbers do not give a full answer, they at least suggest that part of MMAD's strong rise in regional protest may be due to real changes in regional protest patterns rather than simply the shifting attention of international journalists.

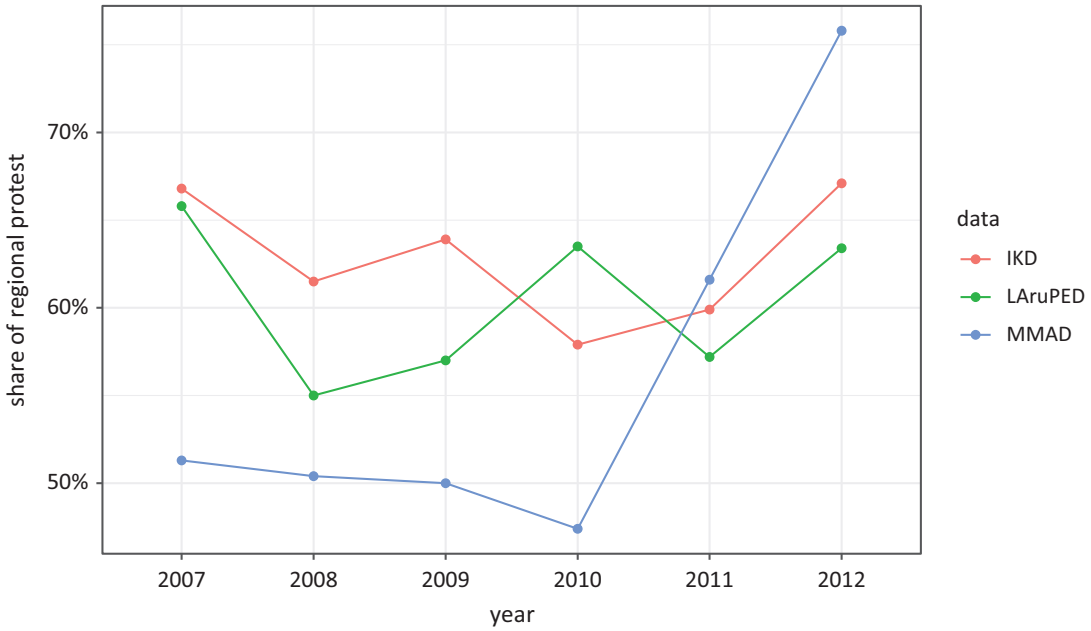

Figure 2. Share of protest events outside Moscow and St. Petersburg over time, by data source (March 2007–March 2012).

5.3. Coverage Across Regions

The two previous sections have shown broadly similar trends across the two activist-based data sets. Regarding regional protest, MMAD did not stray too far from the other two data sets, albeit on a different baseline and except for the high phase of the FFE protests. This leads to the question of whether these trends are only comparable for the country as a whole or whether they hold true across specific regions. Table 2 lists the regions with the highest event counts in the jointly covered period (out of 83 covered regions in total). Unsurprisingly, Moscow and St. Petersburg come out on top in all three data sets. Sverdlovsk, Kaliningrad, and Samara are also among the top 10. But there is also considerable variation: In LAruPED, Penza ranks fifth but is placed much lower in MMAD (rank 43) and IKD (rank 34). Similarly, Dagestan in the North Caucasus is surprisingly highly ranked in MMAD, which is mirrored neither in the LAruPED nor in the IKD data (ranks 35 and 38, respectively).

Do the data show entirely different pictures, or are these dissimilarities exceptions? To answer this question, I compute Spearman's rho of all of the 83 regions covered for all three pairs of data sets. Spearman's rho is better suited for this task than Pearson's r, because it compares the ranks of the regions rather than the absolute values, limiting outliers to the values of their rank. Figure 3 shows that the two activist-based data sets correlate strongly at *rho* = 0.79 ($p < .001$). The regions are not ranked identically—there is considerable variability as to how many events are covered per region—but they are usually in similar sections of the rank order. This indicates that broadly speaking, LAruPED and IKD document similar things. This is noteworthy and encouraging given the different political projects behind them. The same cannot be said for the correlations between MMAD and either of the other two, where Spearman's rho is at only moderate levels (0.48 and 0.59). This suggests that, while bearing some relation to the activist data, international media paint a rather different picture (certainly in part because of MMAD's inclusion threshold).

Having demonstrated that the MMAD data appear to cover protest in Russia's regions differently, I explore one of these differences to the activist-based data sets in greater detail. Table 2 shows Dagestan and North Ossetia, two regions of the North Caucasus, to be in MMAD's top 10. Both are ranked far lower in LAruPED and IKD. Table 3 compares the six core ethnic republics of the Russian North Caucasus—Chechnya, Dagestan,

Table 2. Regions with the highest number of events per data source (March 16, 2007–March 5, 2012).

MMAD			LAruPED			IKD		
Region	# of events	% of total	Region	# of events	% of total	Region	# of events	% of total
Moscow City	346	30.0%	Moscow City	1,288	29.6%	Moscow City	1,795	32.5%
St. Petersburg	152	13.2%	St. Petersburg	484	11.1%	St. Petersburg	309	5.6%
Dagestan	54	4.7%	Samara	176	4.0%	Leningrad	232	4.2%
Primorie	40	3.5%	Moscow Oblast	130	3.0%	Novosibirsk	166	3.0%
Sverdlovsk	37	3.2%	Penza	115	2.6%	Sverdlovsk	144	2.6%
Novosibirsk	27	2.3%	Sverdlovsk	104	2.4%	Samara	138	2.5%
Kaliningrad	23	2.0%	Kaliningrad	99	2.3%	Irkutsk	131	2.4%
Samara	20	1.7%	Voronezh	99	2.3%	Moscow Oblast	130	2.4%
Bashkortostan	19	1.6%	Primorie	96	2.2%	Kaliningrad	125	2.3%
North Ossetia	19	1.6%	Kirov	89	2.0%	Chelyabinsk	122	2.2%

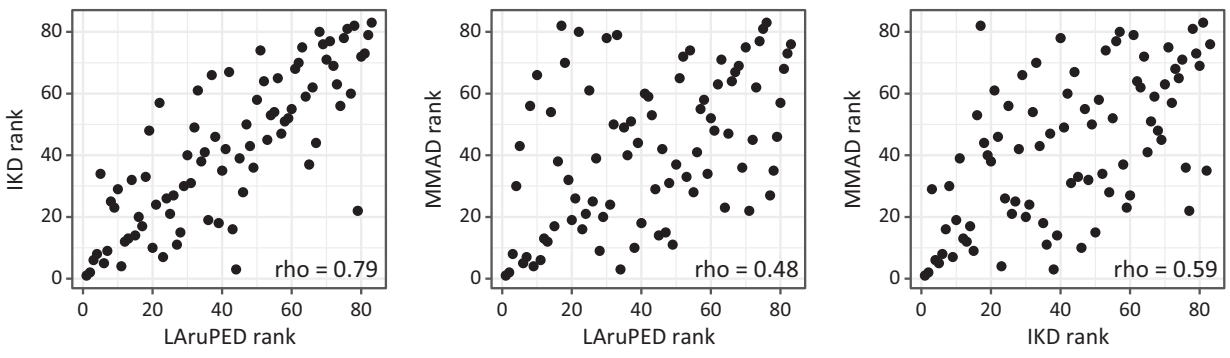

Figure 3. Rank correlations of the number of covered protest events per region, by combination of data sets (March 2007–March 2012).

Table 3. Protest events in North Caucasus by data source and year.

	2007	2008	2009	2010	2011	2012
MMAD	25 (31.6%)	31 (51.7%)	10 (14.3%)	11 (11.8%)	32 (12.4%)	5 (5.3%)
LAruPED	18 (4.6%)	16 (3.0%)	17 (2.7%)	21 (4.0%)	8 (2.0%)	3 (3.5%)
IKD	14 (2.7%)	19 (2.9%)	16 (1.8%)	17 (2.5%)	4 (0.7%)	1 (1.0%)

Notes: Cells show absolute numbers of reported events in the North Caucasus and the share among all regional protest (in brackets). North Caucasus defined as Chechnya, Dagestan, Ingushetia, Kabardino-Balkaria, Karachaevo-Cherkessia, and North Ossetia.

Ingushetia, Kabardino-Balkaria, Karachaevo-Cherkessia, and North Ossetia—demonstrating that this is a systematic difference. MMAD draws a much larger share of its regional protest events from that macro-region, particularly before 2011. Moreover, even though it covers far fewer regional protest events overall, in the Caucasus MMAD has higher *absolute* numbers in several of the years covered.

This finding allows two conclusions: On the one hand, the strong oversampling of the Caucasus compared to other regions indicates a bias resulting from reliance on English-speaking news, as the authors of MMAD note themselves. The North Caucasus being a region with a long history of instability and conflict, it might be that international news agencies have a particular focus on this macro-region. On the other hand, the comparison of absolute numbers also shows that the activist-based data sources under-report on the North Caucasus. Given the highly repressive context, where activist groups such as Other Russia (the source for LAruPED) have a poor standing, this is hardly surprising. But it points to an insight of broader relevance. Where the freedom of the press is limited, activist groups may be an important alternative source for protest data, but where freedoms are curtailed to such an extent that activist groups cannot associate freely, accurate protest coverage depends on either international journalists (who might enjoy a somewhat higher level of protection) or on local sources that are less formalized—and thus less easily targeted by repression—than the activist groups behind LAruPED and IKD. A look into the sources of the MMAD for the North Caucasus reveals that most come from the BBC World Monitoring service that translates local sources, making the latter the more likely explanation. At any rate, a hypothesis for further research could hence be that regime features may play an important role in the decision of which data to use for which questions.

6. Empirical Analysis II: Comparing Event Coverage

The previous section has looked at protest from a bird's eye perspective. I now turn to a brief comparison of the specific coverage of protest in two regions, comparing the activist-based data sets to Semenov's CPR data from Perm and Tyumen. Since Stage 1 has suggested that MMAD underreports regional protest in "normal times," i.e., in the absence of a cross-nationally diffusing protest wave, MMAD is excluded from this comparison.

6.1. Overlap in Event Coverage

A pressing question when comparing specific coverage is to what extent the data sets cover the same events. I approximate this by computing the overlap of the three data sets using the date as an indicator. When two data sets have an entry on the same date in the same region, these are counted as the same event. Certainly, this is a crude method. It would be preferable to use additional characteristics such as precise location or action form, which are, however, not given in the data. The Supplementary File contains at least a partial validity check of overlap between LAruPED and CPR; using the available information on specific events, I show that the date-based method is a useful approximation, but somewhat overestimates the overlap. This is a clear limitation. However, the crude method can still produce valid insights if the limitation is used productively: Calculating overlap only based on the date likely produces false positives, i.e., events that are classified as the same although they are not. At the same time, the method is unlikely to produce false negatives, i.e., events that are coded as different although they, in reality, are the same. Provided that the dates are assigned correctly, this procedure constitutes a most-likely case to detect convergence—meaning that any indication of difference under this scenario would likely translate into even larger differences if the events were identified in a more sophisticated way.

Figure 4 computes the overlap between the three data sets for both regions separately. For each region, it shows what portion of the overall date-region combinations (here treated as the same events) are covered only by either of the three data sets, by two of the three data sets, or by all of them. These Venn diagrams, therefore, visually illustrate the overlap between the data sets. The colouring, moreover, underscores the percentages displayed in each of the sections: White means a low share of protest events, red means a high share.

As the diagrams show, CPR covers by far the most events in both regions, proving that neither IKD nor LAruPED cover the full protest landscape. In Perm, 31% of all events are uniquely covered by CPR, in Tyumen that share stands at 65% (shown in the top section of the top circles in each diagram). This suggests that the Tyumen data, which contain official sources, are even more comprehensive than the Perm data that are based on media alone.

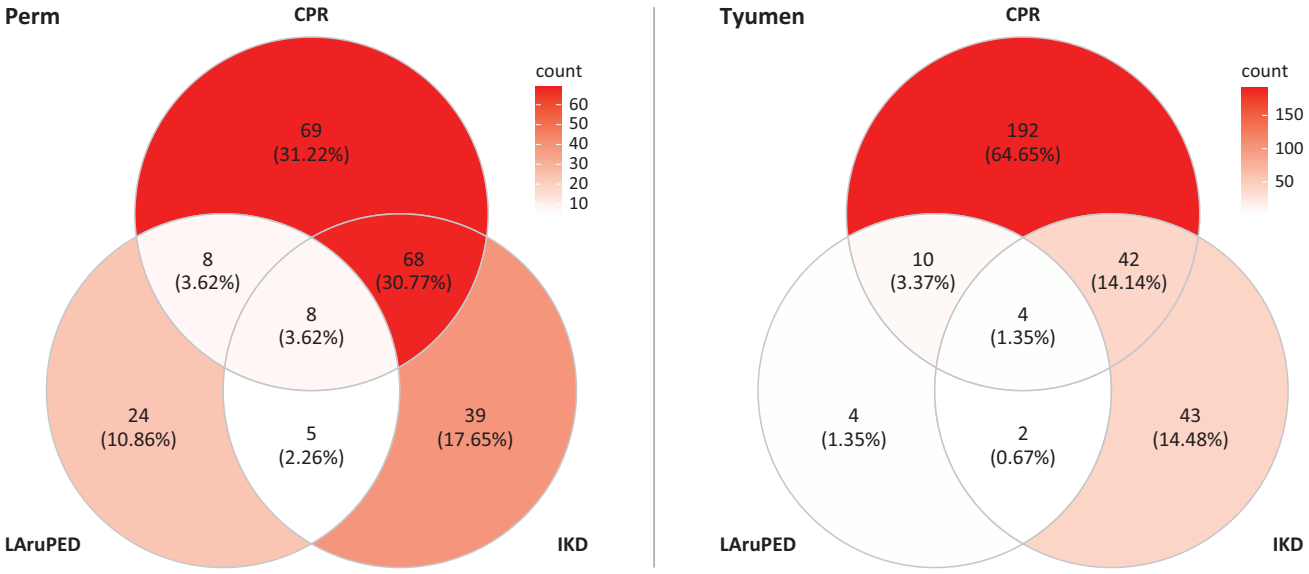

Figure 4. Overlap of three data sets, by region: Left panel shows Perm (March 16, 2007, through March 5, 2012); right panel shows Tyumen (January 1, 2008, through to March 5, 2012).

These findings confirm H4 and H5: In line with H4, CPR has a higher absolute number of covered events than LAruPED or IKD; in line with H5, when comparing the two cities, the CPR data from Tyumen have a larger share of unique events than they have in Perm, underlining the importance of diversifying the source base when attempting to approach full coverage. However, contrary to the assumption behind the two hypotheses, the figures show that CPR does not constitute full coverage either. Between 15% (Tyumen) and 30% (Perm) of all events are not covered by CPR. Therefore, even though the "blanketing strategy" generates more data than the activist-based approach, dissident websites appear to have a unique added value. The lower share in Tyumen suggests that this can be reduced by diversifying source types, but it cannot be eliminated.

Finally, the two different activist data sets appear to have quite different foci, as the overlap between LAruPED and IKD is rather small. Given that, in the country-wide perspective, the two have shown similar trends, the regions of Perm and Tyumen might be outliers. However, the observed divergence suggests that all researchers who are interested in specific regional protest events are well-advised to approach any single data set with caution.

6.2. Thematic Coverage

As the last step, I compare thematic protest coverage in the two regions between CPR and LAruPED. The latter is chosen because of its intuitive and easily adaptable coding scheme (Lankina, 2018), which was applied to CRP (see Supplementary Files). Table 4 displays protest in the two regions for each of LAruPED's six thematic categories as a share of the total number of coded topics. For Perm, the two data sets provide strikingly similar distributions. Except for economic protests (mostly against low salaries or wage arrears), the difference between the two data sets is marginal. Contrary to H6, LAruPED does not oversample political protest relative to a media-based approach. These results are especially noteworthy considering the previous analysis that showed only a small overlap between the two data sets on specific events. The results thus suggest that more efficient methods like LAruPED's *can* come to similar conclusions as the resource-intensive mining of local sources.

The numbers on Tyumen show, however, that this is not a given. Here, LAruPED clearly overreports political and underreports social and environmental protest in comparison to CPR. If these findings reflect more than a coincidence, one could tentatively conclude that the

Table 4. Topics of protest as a share of all, by region and data source.

Region	Data	political	economic	social	cultural	legal	environm.	# of events
Perm	LAruPED	27.7%	23.4%	29.8%	0.0%	8.5%	10.6%	45
Perm	CPR	27.6%	17.6%	31.8%	0.0%	12.9%	10.0%	167
Tyumen	LAruPED	52.4%	9.5%	23.8%	9.5%	4.8%	0.0%	20
Tyumen	CPR	39.2%	7.2%	30.1%	3.3%	11.0%	9.1%	295

Note: Perm covered March 16, 2007, through March 5, 2012; Tyumen covered January 1, 2008, through March 5, 2012.

activist nature of LAruPED seems to drive over-coverage of political events where the absolute number of events is low. In Perm, LAruPED contains 45 events, more than twice the number that it covers in Tyumen. Since this hypothesis cuts to the core of a tradeoff between resources and accuracy in protest event analysis data, it should be investigated more systematically in further research.

7. Discussion and Conclusion

This article gives an overview of four protest event data sets that use international media, dissident websites, and diverse local sources (including news reporting and official accounts). It pursued the question of whether the pictures that each data set paints of protest in Russia differ, and if so, in what way. One answer is bad news: what one learns about protest in Russia depends in part on the chosen data. The good news, however, is that these different pictures at least partially overlap, providing confidence that the data constitute more than statistical noise, and that, when carefully matching one's research question with the appropriate data, valid inferences can be drawn from all of them. To conclude, I differentiate three levels of analysis to briefly discuss the strengths and weaknesses that each type of data set provides with regard to specific research aims.

On the national level, protest event analysis can be used to identify "extraordinary times," periods of high protest activity that are the most likely to provoke a reaction from political authorities and are thus important when studying regime dynamics. The MMAD data may be best suited for this research interest: They identify the fewest events overall, but they clearly mark the protest wave of 2011–2012 (including its regional component), which arguably had the greatest effect on Russian politics of all protest periods in post-Soviet history (Dollbaum, 2020b; Greene & Robertson, 2019). The fact that MMAD spikes in 2011–2012 and thus clearly delineates the FFE period is, as comparison with LAruPED has shown, in part the result of more events crossing the 25-participant threshold. It might, in addition, also be driven by the greater attention of international news media. These two factors, then, make the MMAD data quite efficient in identifying periods of particular protest intensity that are most likely to have political consequences.

If one is interested in protest dynamics on a cross-regional level rather than on its overall effects on the regime as a whole (Gabowitsch, 2016), then the efficient MMAD approach distorts the picture. It might cover single, highly repressive regions better (see the results on the North Caucasus), but the general lesson is that, for studying regional trends, regional data are necessary. Here, the encouraging news is that LAruPED and IKD converge to a high extent, even though they do not cover the same specific events.

Finally, on a case study level, even the activist-based data strongly underreport absolute numbers. If single protest events are of interest, the analysis suggests that one cannot easily circumvent the cumbersome blanketing strategy which should, however, include activist data to reap their unique benefits. If, by contrast, the goal is not the single event but the distribution of protest topics, the comparison of Perm and Tyumen suggested that activist data may give a relatively accurate account—but only if they cross some threshold of absolute coverage. This last point, however, is an inductively generated hypothesis that needs to be investigated in further research. Table 5 summarizes the strengths and weaknesses of the different data set types.

Overall, the comparison gives reason for both confidence and caution. The most important insight is not that any data source outperforms another, but that researchers should invest time in matching their research goals to the data they use in pursuing them. This, of course, presupposes that different data sources exist to choose from.

This, finally, opens the question of how the findings—and conclusions drawn from them—travel to other contexts. In many other non-democratic regimes, the source base will be less generous than it is in Russia—for instance, because of a lack of well-established opposition

Table 5. Advantages and disadvantages of data sets compared in the analysis.

Data set	Sources	Advantages	Disadvantages	Level of analysis for best use
MMAD	International news reports	Identifies major protest episodes Available cross-nationally	Distorts cross-regional dynamics (especially of small protest events) Underreports absolute numbers	National/cross-national
LAruPED IKD	Activist-based	Data sets converge on regional dynamics Available cross-regionally	Underreport absolute numbers	Subnational comparative
CPR	Multiple local sources	Fairly comprehensive (but still not full coverage)	Resource intensive	Subnational case-study

projects like namarsh.ru or IKD or because of tighter restrictions on the online sphere. Under such conditions, it will not be possible to use different data sources for different research purposes. Nonetheless, findings from this study may inform the methodological discussion and the application of protest event analysis more broadly. Beyond providing empirical support for the requirement to match research question and data source, the findings suggest particular ways in which different types of data sources systematically differ in the way they cover protest. For instance, if the conclusions on the MMAD data are valid, then, in countries where data based on international sources are the only data available, it should be possible to identify periods of high protest intensity that are likely to trigger political responses (repression, concessions, etc.). Conversely, the analysis has shown that MMAD will be of less use when studying the subnational dynamics of protest. Moreover, the two activist-based data sets give a relatively similar picture of protest dynamics across regions even though they come from quite different activist groups, reducing fears of strong distortions introduced by such sources. Finally, however, the data show that no single source is sufficiently well-placed to provide a complete picture of protest in the studied period. This analysis, then, serves as a general reminder of the limits of our inferential capacity.

Acknowledgments

This publication was produced as part of the research project Comparing Protest Actions in Soviet and Post-Soviet Spaces—Part 2, which is organized by the Research Centre for East European Studies at the University of Bremen with financial support from the Volkswagen Foundation.

Conflict of Interests

The author declares no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Beissinger, M. R. (2002). *Nationalist mobilization and the collapse of the Soviet state*. Cambridge University Press.

Bremer, B., Hutter, S., & Kriesi, H. (2020). Dynamics of protest and electoral politics in the Great Recession. *European Journal of Political Research*, *59*(4), 842–866. https://doi.org/10.1111/1475-6765.12375

Dollbaum, J. M. (2020a). When does diffusing protest lead to local organization building? Evidence from a comparative subnational study of Russia's "For Fair Elections" movement. *Perspectives on Politics*. Advance online publication. https://doi.org/10.1017/S1537592720002443

Dollbaum, J. M. (2020b). Protest trajectories in electoral authoritarianism: From Russia's "For Fair Elections" movement to Alexei Navalny's presidential campaign. *Post-Soviet Affairs*, *36*(3), 192–210. https://doi.org/10.1080/1060586X.2020.1750275

Downs, A. (1972). Up and down with ecology: The "issue-attention cycle." *National Affairs*, *48*, 38–50. https://www.nationalaffairs.com/storage/app/uploads/public/58e/1a4/b56/58e1a4b56d25f917699992.pdf

Earl, J., Martin, A., McCarthy, J. D., & Soule, S. A. (2004). The use of newspaper data in the study of collective action. *Annual Review of Sociology*, *30*(1), 65–80. https://doi.org/10.1146/annurev.soc.30.012703.110603

Gabowitsch, M. (2016). *Protest in Putin's Russia*. Polity Press.

Geddes, B., Wright, J., & Frantz, E. (2014). Autocratic breakdown and regime transitions: A new data set. *Perspectives on Politics*, *12*(2), 313–331. https://doi.org/10.1017/S1537592714000851

Gladun, A. (2020). Protesting that is fit to be published: Issue attention cycle and nationalist bias in coverage of protests in Ukraine after Maidan. *Post-Soviet Affairs*, *36*(3), 246–267. https://doi.org/10.1080/1060586X.2020.1753428

Greene, S. A., & Robertson, G. B. (2019). *Putin v. the people: The perilous politics of a divided Russia*. Yale University Press.

Hanna, A. (2017). *MPEDS: Automating the generation of protest event data*. SocArXiv. https://doi.org/10.31235/osf.io/xuqmv

Hellmeier, S., Weidmann, N. B., & Rød, E. G. (2018). In the spotlight: Analyzing sequential attention effects in protest reporting. political communication, *35*(4), 587–611. https://doi.org/10.1080/10584609.2018.1452811

Herkenrath, M., & Knoll, A. (2011). Protest events in international press coverage: An empirical critique of cross-national conflict databases. *International Journal of Comparative Sociology*, *52*(3), 163–180. https://doi.org/10.1177/0020715211405417

Hutter, S. (2014). Protest event analysis and its offspring. In D. della Porta (Ed.), *Methodological practices in social movement research* (pp. 335–367). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198719571.001.0001

Keremoglu, E., Hellmeier, S., & Weidmann, N. B. (2020). *Coding instructions for the mass mobilization in autocracies database, version 3.0*. Mass mobilization in autocracies database. https://mmadatabase.org/about/documentation

Kriesi, H., Hutter, S., & Bojar, A. (2019). Contentious episode analysis. *Mobilization: An International Quarterly*, *24*(3), 251–273. https://doi.org/

10.17813/1086-671X-24-3-251
Kriesi, H., Koopmans, R., & Duyvendak, J. W. (Eds.). (1995). *New social movements in Western Europe: A comparative analysis*. University of Minnesota Press.
Lankina, T. (2015). The dynamics of regional and national contentious politics in Russia: Evidence from a new dataset. *Problems of Post-Communism*, *62*(1), 26–44. https://doi.org/10.1080/10758216.2015.1002329
Lankina, T. (2018). *Lankina Russian protest event dataset* [Data set]. http://eprints.lse.ac.uk/90298
Lankina, T., & Tertytchnaya, K. (2020). Protest in electoral autocracies: A new dataset. *Post-Soviet Affairs*, *36*(1), 1–17. https://doi.org/10.1080/1060586X.2019.1656039
Lankina, T., & Voznaya, A. (2015). New data on protest trends in Russia's regions. *Europe-Asia Studies*, *67*(2), 327–342. https://doi.org/10.1080/09668136.2014.1002696
Levitsky, S., & Way, L. A. (2010). *Competitive authoritarianism: Hybrid regimes after the Cold War*. Cambridge University Press.
McCarthy, J., Titarenko, L., McPhail, C., Rafail, P., & Augustyn, B. (2008). Assessing stability in the patterns of selection bias in newspaper coverage of protest during the transition from communism in Belarus. *Mobilization: An International Quarterly*, *13*(2), 127–146.
McFaul, M., & Stoner-Weiss, K. (2008). The Myth of the authoritarian model: How Putin's crackdown holds Russia back essay. *Foreign Affairs*, *1*, 68–84.
Oliver, P. E., & Maney, G. M. (2000). Political processes and local newspaper coverage of protest events: From selection bias to triadic interactions. *American Journal of Sociology*, *106*(2), 463–505. https://doi.org/10.1086/316964
Petrov, N., & Titkov, A. (2013). *Rejting demokratichnosti regionov Moskovskogo Centra Karnegi: 10 let v stroju* [The Carnegie Moscow Center's Rating or Regional Democracy: 10 years in the making]. Carnegie Endowment for International Peace.
Reuter, O. J., & Robertson, G. B. (2015). Legislatures, cooptation, and social protest in contemporary authoritarian regimes. *The Journal of Politics*, *77*(1), 235–248. https://doi.org/10.1086/678390
Robertson, G. B. (2013). Protesting Putinism: The election protests of 2011–2012 in broader perspective. *Problems of Post-Communism*, *60*(2), 11–23. https://doi.org/10.2753/PPC1075-8216600202
Semenov, A. (2017). From economic to political crisis? Dynamics of contention in Russian regions (2008-2012). *Österreichische Zeitschrift Für Politikwissenschaft*, *45*(4). https://doi.org/10.15203/ozp.1107.vol45iss4
Tarrow, S. G. (1993). Cycles of collective action: Between moments of madness and the repertoire of contention. *Social Science History*, *17*(2), 281–307. https://doi.org/10.2307/1171283
Tilly, C., Tilly, L., & Tilly, R. (1975). *The rebellious century*. Harvard University Press.
Weidmann, N. B., & Rød, E. G. (2019). *The internet and political protest in autocracies*. Oxford University Press.
Wouters, R. (2013). From the street to the screen: Characteristics of protest events as determinants of television news coverage. *Mobilization: An International Quarterly*, *18*(1), 83–105. https://doi.org/10.17813/maiq.18.1.y6067731j4844067
Wüest, B., & Lorenzini, J. (2020). External validation of protest event analysis. In H. Kriesi, J. Lorenzini, B. Wüest, & S. Hausermann (Eds.), *Contention in times of crisis: Recession and political protest in thirty European countries* (pp. 49–78). Cambridge University Press.
Zhang, H., & Pan, J. (2019a). CASM: A deep-learning approach for identifying collective action events with text and image data from social media. *Sociological Methodology*, *49*(1), 1–57. https://doi.org/10.1177/0081175019860244
Zhang, H., & Pan, J. (2019b). The challenges of "more data" for protest event analysis. *Sociological Methodology*, *49*(1), 76–82. https://doi.org/10.1177/0081175019866425

About the Author

Jan Matti Dollbaum is a postdoctoral researcher at the University of Bremen. He studies social movements and their interaction with institutional politics across different regime types. His work has appeared in *Perspectives on Politics*, *Post-Soviet Affairs*, and *Social Movement Studies*, among others. Together with Ben Noble and Morvan Lallouet he is the author of *Navalny: Putin's Nemesis, Russia's Future?* (Hurst Publishers/Oxford University Press).

Media and Communication (ISSN: 2183-2439)

Media and Communication is an international open access journal dedicated to a wide variety of basic and applied research in communication and its related fields. It aims at providing a research forum on the social and cultural relevance of media and communication processes.

www.cogitatiopress.com/mediaandcommunication

COGITATIO