Shaping Online News Recommendations in Russia: The Yandex.News Controversies

Daucé, Françoise; Loveluck, Benjamin

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

gesis
Leibniz-Institut
für Sozialwissenschaften

Mitglied der
Leibniz-Gemeinschaft

Shaping Online News Recommendations in Russia: The Yandex.News Controversies[1]

By Françoise Daucé (School for Advanced Studies in the Social Sciences (EHESS)) and Benjamin Loveluck (i3-SES)

Abstract

In Russia, numerous controversies have arisen since 2012 around the political role of the aggregator Yandex.news in prioritizing media news. Through its algorithm, this service is suspected of contributing to the decline of information pluralism for political purposes. These suspicions have only grown with the start of the war in Ukraine.

Where Google Does Not Dominate

Russia is among the few countries in the world where Google does not dominate the online search industry. In 2020, the Russian-language equivalent, Yandex, held just under half the market share (about 45 percent).

Yandex has long benefitted from a certain degree of autonomy, and its founders have even, at different moments, expressed political disagreement with the Kremlin. However, as a national economic champion and a key player in the organization of information, it has found itself under tight scrutiny. This has been particularly true since the 2011–2012 protests against electoral fraud and the 2014 annexation of Crimea, which also represented turning points for Russia due to the increased control exerted over the media, Internet, and civil society (Oates, 2013; Soldatov and Borogan, 2015; Wijermars and Lehtisaari, 2020).

A case in point is the Yandex.Novosti ("Yandex. News") aggregator—the Russian equivalent of Google News, launched in 2004—which is the focus of this article. When they first appeared, search engines and recommendation systems such as aggregators were designed as tools that would make the diversity of content on the Web more manageable. As a vast body of research has shown, however, these platforms occupy a strategic place and have become key intermediaries in channeling information to end users *qua* citizens. Thus, they wield a form of power in shaping users' perception of social reality that scholars, policymakers and civil society alike are still in the process of defining. With the start of Russia's war against Ukraine on February 24, 2022, the role of Yandex.News in controlling the media agenda in Russia has become an even more crucial issue.

Algorithmic Gatekeeping in Digital Media Ecosystems

The Yandex.News aggregator can be described as an *automated news recommender system*. The best-known example of such a service is the Google News aggregator, which was first launched in 2002 and taken out of beta in 2006 (Bharat, 2006). Initially, the service aimed to provide a broad overview of trending news by presenting users with "clusters" of related articles. As of 2021, the service indexed tens of thousands of news websites around the world and was woven into Google's main web search service.

Google and Yandex alike have generally presented their services as "neutral," but such claims to objectivity have been criticized for various reasons. For the past decade, because of their increasingly powerful personalization features, some of the main Web services—and particularly Google's search engines—have been suspected of entrapping users in "filter bubbles" and "echo chambers" (Pariser, 2011; Bozdag, 2013). By making users oblivious to certain types of information or to alternative perspectives, and by sometimes reinforcing existing prejudices or biases, these services arguably undermine the public sphere. Search algorithms and automated recommender systems have also been criticized for promoting outrage and conspiracy theories, with the YouTube recommendation algorithm, for instance, being presented as "the great radicalizer" (Tufekci, 2018).

However, the reality of these phenomena is difficult to assess precisely (Flaxman et al., 2016; Bruns, 2019), particularly in the case of search engines, which have also been shown to increase information diversity (see Fletcher and Nielsen, 2018). The algorithms deployed by these platforms can therefore be perceived as an "invisible hand," deciding which topics will be singled out as relevant and which news outlets will be pushed to the forefront according to sometimes unfathomable criteria—profoundly affecting the nature of journalism in the process, as professionals adjust the form and nature of their published content in line with these constraints

---

1   A longer version of this article has been published online: Daucé, Françoise, and Benjamin Loveluck. "Codes of conduct for algorithmic news recommendation: the Yandex. News controversy in Russia." First Monday (2021).

(Brake, 2017; Christin, 2020). In the Russian political context, the issue raised by the Yandex.News aggregator is acute: could it be manipulated for political reasons, either through direct interference with the results or by fooling its algorithm?

Yandex.News as Political Controversy

Russia's political leadership has targeted Yandex.News through various policies and legal initiatives since 2014. Yandex.News presents a selection of topics and articles that purport to reflect the themes most widely covered by the media at any given moment. To do so, it processes the information published by a range of (mainly Russian) online media. Yandex.News was launched in 2004 and was initially a pilot project led by a team of computer scientists and linguists who had been hired to develop named-entity recognition and extraction in the news. The Yandex.News team claims that the algorithm works in the absence of human intervention. News from partners is gathered into topics through the algorithm's clustering process, which analyzes keywords and facts using three main criteria: citation rate, recency, and informativity. A Top 5 of its aggregation results is always visible on the Russian version of the Yandex homepage, just above the search box. In 2017, according to Grigori Bakunov, Yandex Technical Director, "The daily audience of the five news items that appear on the Yandex homepage is the same as the homepage—approximately 20 million people, depending on the day. Six million visit the Yandex.News page daily."

However, the controversies that arose after 2012 put an end to public belief in the objectivity of the aggregator. That year was a decisive one for freedom of expression in Russia and a "watershed moment" for Internet regulation (Lonkila et al., 2020). Control over the public sphere increased again in 2014 during the conflict with Ukraine and the occupation of Crimea. Yandex.News, in particular, found itself at the heart of a political controversy after being accused of partiality by the authorities for providing visibility to information that did not align with the official narrative. The site Pravda.ru wondered whether "Yandex lights a 'Maidan' in Russia?" (referring to the protests in Kiev that led to regime change in Ukraine).[2] The newspaper was outraged by the headlines chosen by the news aggregator and claimed that legal regulation of its activity was required. This led to the adoption in 2016 of a law on news aggregators that was designed to extend control of the media to such intermediaries and specifically targeted Yandex.News.[3] Those news aggregators that received over one million daily visitors became legally responsible for any content published in their results (and at risk of heavy fines in the event of violations), unless the selected media had been officially registered with Roskomnadzor. The law went into effect on January 1, 2017, whereupon all non-registered media (including dissenting voices such as Mediazona), as well as all foreign media (such as the BBC in Russian and exiled media such as Meduza), disappeared from both the Top 5 results presented on the Yandex homepage and Yandex.News. In sum, the aggregator may claim to be neutral and objective, but on the one hand, the authorities denounce its propensity to relay discontent and destabilize the political situation, while on the other hand, journalists, web professionals, and activists underline that its institutional framing requires it to promote a "loyal" agenda.

Its shortcomings have been made clear since Russia's invasion of Ukraine, as the last remaining independent media have gradually been shut down. In a post published on his Facebook page three days after the beginning of the war, Lev Gershenzon, former head of Yandex.News, stated:

> Now every day Russia's war against Ukraine is possible because there are no mass anti-war demonstrations in Russian cities. And they don't happen not only because of the danger of reprisals to those who do come out (huge admiration to all who do come out), but mainly because the vast majority of the population is unaware that Russian troops are in their fourth day of full-scale warfare. Leading this ignorance, along with television, is Yandex—a website and apps with a news bloc, 5 news, "on the home page." This news gets straight to people precisely because they do not come for it, but for some other reason: to find a product or the address of a pharmacy, to see the dollar exchange rate or the weather, etc. We once articulated that the task of this unit was to find out "if anything is wrong." So now it says: "no, there is no problem." […] Every hour and day that it works the way it does now is an endorsement of the war.[4]

The Yandex Rankings as a Gateway to the Algorithm and Its Transformations

Though it is difficult to investigate the algorithm itself, one can look at the output that the aggregator displays. During the month of June 2020, we conducted a quantitative analysis of the news selected by Yandex.News and presented as part of the Top 5 on the Yandex homepage.

---

2 "Yandex 'razzhigaet' Majdan v Rossii?", Pravda, at http://www.pravda.ru/topic/yandex-617/.
3 Federal Law № FZ-208, 23 June 2016.
4 Lev Gershenzon Facebook page, February 27, 2022.

We carried out a systematic scraping of news: between June 1 and June 30, 2020, we automatically collected the Yandex.News rankings every two hours and listed a total of 3,011 references.[5] It appeared that, during this period, only 14 media outlets were cited in the Top 5—an extremely narrow sample considering the more than 7,000 sources listed in the Yandex.News database. We then extended the scraping to the period June–December 2020 and obtained the same results, with the same 14 media appearing in the Top 5 over this period. The data provide striking evidence of the concentration of information on Yandex.News with a few large media players: public press agencies, state-funded media, leading newspapers, and mainstream online publications (RIA Novosti, Gazeta.ru, Izvestia, RBK, Lenta.ru, RT in Russian, Kommersant, Regnum, Rossiiskaia Gazeta, TASS, Vesti.ru, Vedomosti, BFM.ru, and Interfax).

The over-representation of specific news publishers has also been demonstrated in the case of Google News (Schroeder and Kralemann, 2005; Haim et al., 2018), but not to such an extent. Our results from Yandex.News feature a much narrower range of publications than the findings of Nechushtai and Lewis (2019) in the case of Google News in the US, for instance. Although 14 outlets likewise dominated that aggregator, a long tail of other publications also figured in the results. Moreover, even if nuances can be detected between the 14 major media that dominate Yandex.News in terms of their editorial line, it is evident that in 2020, "officially sanctioned" media reached Yandex's heights more easily. Indeed, most of the 14 selected outlets are related to the Kremlin: they are either funded by the state directly or are privately owned by "loyalist" figures or entities and thus indirectly "managed" by the authorities.

The recent history of Yandex.News in Russia highlights how platform regulation can be leveraged to set up a form of "governance by algorithms" of the media and the public sphere. Initially presented as a technical means to "objectively" assess the diversity of online content, the aggregator sparked techno-political controversy in the 2010s: it was criticized by the authorities for promoting "unpatriotic" or "fake" news, while journalists, web professionals, and end users increasingly suspected that *inconvenient* truths would find it difficult to reach its top rankings. The adoption in 2016 of a law on news aggregators, which allowed only officially "registered" sources to be displayed by the service, clearly reflected an intention on the part of the authorities to domesticate the platform in order to limit the visibility of protests and discontent in the public sphere. This regulation took place in a complex digital ecosystem that articulates different levels of gatekeeping, including Yandex.News and other platforms, the telecommunications watchdog Roskomnadzor, as well as media outlets and journalists.

Yandex as a news recommender system abides by both legal and technical "codes of conduct" that help ensure that the information it promotes and amplifies remains in check. Although no outright censorship has yet been demonstrated at the level of Yandex.News, the aggregator appears to be an important cog in the machine of tightening control exerted by the authorities over the overall Russian media ecosystem. Until recently, however, governance by algorithms has remained imperfect and taken place in an intricate technical, political, legal, and economic context where national and international platforms have coexisted and competed.

Journalists and publishers could seek alternative channels to distribute information, relying on social media such as Telegram or Twitter. It remains to be seen how far this will still be possible as the war unfolds and the space for critical voices diminishes. Up until the beginning of the war, the Russian authorities justified their efforts to control the media agenda and to reassert their sovereignty over the public sphere by denouncing information framed as "unpatriotic," "fake" or otherwise problematic. Today, any dissenting views are being quelled, and the role of Yandex.News is at the heart of political concerns about the use of algorithms for warmongering.

About the Authors
Françoise Daucé is Professor at the School for Advanced Studies in the Social Sciences (EHESS), Paris, France and director of the Center for Russian, Caucasian and East-European Studies (CERCEC).

Benjamin Loveluck is Associate Professor at i3-SES, Telecom Paris, France.

References
• Bharat, K. (2006). "And now, News," *Google Official Blog*, 23 January 2006, https://googleblog.blogspot.com/2006/01/and-now-news.html

5　We analyzed the code of the Yandex homepage and found that 10 news references were presented at any given time. We therefore set up a Node.js script to collect these 10 references every two hours: four references occupy places 1 to 4 of the Top 5, while the fifth place is likely occupied by the six other references on a rotating basis. The script uses two main Node libraries: Puppeteer for scraping and Mongoose for database registration. After manually analyzing the html code of the homepage and several other pages of the website, we wrote javascript code to scrape the content of the 10 top news (title, date, source name, source url, rank on the homepage). The data was then registered in a MongoDB database using the Mongoose library.

- Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, *15*(3): 209–227.
- Brake, D. R. (2017). The invisible hand of the unaccountable algorithm: how Google, Facebook and other tech companies are changing journalism. In J. Tong & S.-H. Lo (eds.), *Digital Technology and Journalism. An International Comparative Perspective*. Cham: Palgrave Macmillan, pp. 25–46.
- Bruns, A. (2019). *Are Filter Bubbles Real?* Cambridge: Polity Press.
- Christin, A. (2020). *Metrics at Work. Journalism and the Contested Meaning of Algorithms*. Princeton, NJ: Princeton University Press.
- Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, *80*(S1): 298–320.
- Fletcher, R., & Nielsen, R. K. (2018). Automated serendipity: the effect of using search engines on news repertoire balance and diversity. *Digital Journalism*, *6*(8): 976–989.
- Haim, M., Graefe, A., & Brosius, H.-B. (2018). Burst of the filter bubble? Effects of personalization on the diversity of Google News. *Digital Journalism*, *6*(3): 330–343.
- Lonkila, M., Shpakovskay, 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*. Abingdon and New York: Routledge, pp. 17–38.
- Nechushtai, E., & Lewis, S. C. (2019). What kind of news gatekeepers do we want machines to be? Filter bubbles, fragmentation, and the normative dimensions of algorithmic recommendations. *Computers in Human Behavior*, *90*: 298–307.
- Oates, S. (2013). *Revolution Stalled. The Political Limits of the Internet in the Post-Soviet Sphere*. Oxford and New York: Oxford University Press.
- Pariser, E. (2011). *The Filter Bubble. What the Internet Is Hiding from You*. New York: Penguin Press.
- Schroeder, R., & Kralemann, M. (2005). Journalism ex machina—Google News Germany and its news selection processes. *Journalism Studies*, *6*(2): 245–247.
- Soldatov, A., & Borogan, I. (2015). *The Red Web. The Kremlin's Wars on the Internet*. New York: PublicAffairs.
- Tufekci, Z. (2018). YouTube, the great radicalizer. *The New York Times*, 10 October 2018. https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html
- Wijermars, M. & Lehtisaari, K. eds. (2020). *Freedom of Expression in Russia's New Mediasphere*. Abingdon and New York: Routledge.

ABOUT THE RUSSIAN ANALYTICAL DIGEST

Editors: Stephen Aris, Matthias Neumann, Robert Orttung, Jeronim Perović, Heiko Pleines, Hans-Henning Schröder, Aglaya Snetkov

The Russian Analytical Digest is a bi-weekly internet publication jointly produced by the Research Centre for East European Studies [Forschungsstelle Osteuropa] at the University of Bremen (www.forschungsstelle.uni-bremen.de), the Center for Security Studies (CSS) at the Swiss Federal Institute of Technology Zurich (ETH Zurich), the Center for Eastern European Studies at the University of Zurich (http://www.cees.uzh.ch), the Institute for European, Russian and Eurasian Studies at The George Washington University (https://ieres.elliott.gwu.edu), and the German Association for East European Studies (DGO). The Digest draws on contributions to the German-language Russland-Analysen (www.laender-analysen.de/russland), and the CSS analytical network on Russia and Eurasia (www.css.ethz.ch/en/publications/rad.html). The Russian Analytical Digest covers political, economic, and social developments in Russia and its regions, and looks at Russia's role in international relations.

To subscribe or unsubscribe to the Russian Analytical Digest, please visit our web page at http://www.css.ethz.ch/en/publications/rad.html

Research Centre for East European Studies at the University of Bremen
Founded in 1982, the Research Centre for East European Studies (Forschungsstelle Osteuropa) at the University of Bremen is dedicated to the interdisciplinary analysis of socialist and post-socialist developments in the countries of Central and Eastern Europe. The major focus is on the role of dissent, opposition and civil society in their historic, political, sociological and cultural dimensions.
With a unique archive on dissident culture under socialism and with an extensive collection of publications on Central and Eastern Europe, the Research Centre regularly hosts visiting scholars from all over the world.
One of the core missions of the institute is the dissemination of academic knowledge to the interested public. This includes regular e-mail newsletters covering current developments in Central and Eastern Europe.

The Center for Security Studies (CSS) at ETH Zurich
The Center for Security Studies (CSS) at ETH Zurich is a center of competence for Swiss and international security policy. It offers security policy expertise in research, teaching, and consultancy. The CSS promotes understanding of security policy challenges as a contribution to a more peaceful world. Its work is independent, practice-relevant, and based on a sound academic footing.
The CSS combines research and policy consultancy and, as such, functions as a bridge between academia and practice. It trains highly qualified junior researchers and serves as a point of contact and information for the interested public.

The Institute for European, Russian and Eurasian Studies, The Elliott School of International Affairs, The George Washington University
The Institute for European, Russian and Eurasian Studies is home to a Master's program in European and Eurasian Studies, faculty members from political science, history, economics, sociology, anthropology, language and literature, and other fields, visiting scholars from around the world, research associates, graduate student fellows, and a rich assortment of brown bag lunches, seminars, public lectures, and conferences.

The Center for Eastern European Studies (CEES) at the University of Zurich
The Center for Eastern European Studies (CEES) at the University of Zurich is a center of excellence for Russian, Eastern European and Eurasian studies. It offers expertise in research, teaching and consultancy. The CEES is the University's hub for interdisciplinary and contemporary studies of a vast region, comprising the former socialist states of Eastern Europe and the countries of the post-Soviet space. As an independent academic institution, the CEES provides expertise for decision makers in politics and in the field of the economy. It serves as a link between academia and practitioners and as a point of contact and reference for the media and the wider public.