LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-764577

The Russian Model of Internet Control and Its Significance

J. A. Kerr

December 21, 2018

Disclaimer

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

The Russian Model of Internet Control and Its Significance

Jaclyn Kerr, PhD

Abstract: Russia has emerged as an exemplar of an innovative and experimental – though not always completely consistent or successful – alternative approach to information manipulation and control that differs significantly from the more-often discussed Chinese "Great Firewall" system and other approaches with an emphasis on systemic technical censorship. The Russian model relies on a mix of less overt, more plausibly deniable, legalistic, and often non-technical mechanisms to manipulate online information flows, narratives, and framings, to affect and shape public opinion without resort to universal censorship. The government uses surveillance, a panoply of vague laws, the prosecution or censorship of exemplars, proxy actors and hard to track extra-legal pressures, hacking and leaks, and a heavy emphasis on content production and manipulation to influence narratives and shape public opinion. This model for the domestic control of information not only fits with Russia's own political system, but is likely to prove more resonant and easier to emulate across many other countries in which a systematic-censorship approach is not technologically or politically feasible. The learning and experimentation involved in this type of domestic information manipulation also has direct applicability to the use of information operations in international political and military competition. The future of this model will likely depend on continuing innovation, not least on the leveraging of advances in artificial intelligence and big data analysis. If successful, however, this might look very different from the future of information control in China – and have significantly different repercussions for democracies and the international system.

The Internet and new information and communication technologies (ICTs) were once hailed as "liberation technologies" – tools to enable the free flow of information, allowing individual freedoms of expression and organization, and breaking down the last vestiges of authoritarianism. Through the course of the 2000s and early 2010s, while democratic states largely converged on a norm of non-censorship, the most closed authoritarian regimes tended to be early adopters of high-censorship overtly-restrictive approaches to Internet control, adopting such approaches as domestic Internet use levels grew and the required technological solutions became affordable on global or regional markets. But observers questioned the long-term survivability of such adaptations, looking to events such as Iran's Green Movement and the Arab Spring as proof of the vulnerability of non-democratic systems to the new global flows of information and the transformational affordances of the digital technologies.

During this period, "hybrid regimes" – non-democratic regimes that still based their domestic and international legitimacy in part on democratic institutions and rights protections – seemed particularly vulnerable to the sorts of critical discourse and mass protest mobilizations enabled by the new technologies. Fraudulent elections, illegal government actions, corruption, and the inadequate protection of constitutional rights all appeared as potential flashpoints – possibly critical to regime survival but also the potential sources of mass protest around official hypocrisy. These regimes were less quick to adopt systemic censorship or other approaches to Internet control that would overtly violate democratic norms. Such overt actions could further undermine their legitimacy at home and abroad. But they faced increasing pressures to do something – to find alternative approaches to manage the stability risks caused by the increasing use of digital technologies in their societies.

As threats to regime survival in hybrid regime type countries became clear (often following major domestic mass protest mobilizations) these countries began to experiment with alternative mechanisms to rein in the destabilizing influences of the new technologies. These approaches were distinctive from the earlier high-censorship models first adopted in more closed authoritarian regimes, and ultimately more akin to the "low-intensity coercion" approaches these regimes often followed in other areas of domestic political control. This included efforts to utilize democratic legal mechanisms and institutions in combination with pro-regime content production and plausibly deniable forms of disruption to alter online discourse and narratives without recourse to pervasive censorship. Russia emerged as an exemplar of this alternative approach to information control.

A Russian Model of Information Control?

As early as the late 1990s and early 2000s, the Russian government was concerned by the potential destabilizing impact and national security repercussions of information flows within society. In 1995 Russia adopted the "Law on Operational Investigations," giving the FSB authority "to monitor all private communications" of citizens, including electronic communications, and the first "System for Operative Investigative Activities" (or SORM) infrastructure was built – extended in 1998 (SORM-2) to allow monitoring of Internet traffic. Beginning in 1998, Russia submitted nearly-annual resolutions to the United Nations General Assembly concerning "Developments in the field of information and telecommunications in the context of international security," and a 1999 submission to the UN Secretary-General contained a proposed set of "principles in international information security."[1] In these submissions it was clear that the concern related as much to international flows of information *content* as to the growing field of cybersecurity. On September 9th 2000, following the immensely negative media coverage of the Kursk submarine tragedy the previous month, Vladimir Putin (then in his first year of office) signed the new "Information Security Doctrine of the Russian Federation" that had been developed by his Security Council. The document declared formal support for freedom of speech and the media, but it also indicated supposed threats to national security related to the flow of information.

Importantly, despite these moves, during this period the Russian government also took steps towards fuller participation in the global digital economy and to assure their burgeoning domestic Internet industry of this commitment.[2] The Russian Internet (colloquially called "RuNet") developed into a vibrant new space of public discourse, with little or no censorship throughout the 2000s, even as restrictions over mainstream media and civil society tightened. But this is not to say that no effort was made to control the new technology's impact on political stability. This effort increased precipitously following the experience of social-media-fueled mass protest at home.

By the early 2010s, and especially following the 2011-2012 White Ribbon Protest Movement and Vladimir Putin's return to the presidency, Russia emerged as an exemplar of an innovative and experimental – though not always completely consistent or successful – alternative approach to information manipulation and control that differed significantly from the more-often discussed Chinese "Great Firewall" system and other approaches with an emphasis on systemic technical censorship. It has pioneered a distinct model that relies on a mix of less overt, more plausibly deniable, legalistic, and often non-technical mechanisms to manipulate online information flows, narratives, and framings, to affect and shape public opinion without resort to universal censorship. This model for the domestic control of information not only fits with Russia's own domestic political system, but is likely to prove more resonant and easier to emulate across many other countries – including but not limited to other hybrid regimes – in which a systematic-censorship approach is not technologically or politically feasible.

Since 2012, Russia has had a blacklist of legally censored websites. This was a stark change after years in which the Internet was essentially uncensored. But it uses this list parsimoniously, providing legal justifications for each category of restricted content and usually applying these to exemplars rather than systematically. To be clear, pressures on the producers and hosts of controversial online content have

---

[1] The promotion of international standards for information non-aggression became a consistent theme, with Russia also leading blocks of states in efforts. In 2011 and 2015 it collaborated with other countries from the Shanghai Cooperation Organization to submit joint proposals to the UN General Assembly for an "International Code of Conduct for Information Security," for example.

[2] This included, notably, a widely-recalled December 1999 meeting between then-Prime Minister Putin and members of the Russian Internet community in which, under pressure from the assembled bloggers and ISP-directors, Putin rejected a considered plan for more centralized government control over the Internet and promised that they would be consulted before further policy decisions. But some dynamics of consultation continued throughout the 2000s and beyond. During Dmitriy Medvedev's presidency, 2008-2012, Internet entrepreneurship was also avidly promoted as part of his economic modernization program. Medvedev toured Silicon Valley, met with young ICT entrepreneurs, and himself utilizing social media.

increased significantly in Russia in the post-2012 period. But these pressures often take the form of new laws and quasi-democratic processes, financial dealings between companies, or behind-the-scenes (and plausibly deniable requests). A laundry list of new laws have created legal bases for the blocking of a wide variety of content during this period, while also increasing the systematic collection of user data and placing a heavy burden of liability on content intermediaries.

These new laws include, for example:

- The 2013 "Anti-Piracy Law" – This had some similarities to the earlier SOPA/PIPA legislation in the United States which had been rejected due to protests about the impact on Internet intermediaries, but, despite a Russian Internet user protest petition, this one went into force.
- The 2014 "Anti-LGBT Propaganda Law" – This extended the original 2012 Blacklist for the protection of children from child pornography and content related to illegal drugs and suicide, also requiring the blocking of content that could be seen as propaganda for alternative sexual orientations directed at children.
- The 2014 "Law on Pre-Trial Blocking of Websites" – This permitted the immediate blocking of sites deemed to contain "incitement to extremism or riots," and was used to abruptly block several leading oppositional news outlets and blogs at the height of the Crimea Annexation crisis.
- A package of "Anti-Terrorist" laws passed in summer of 2014 – This included the so-called "Blogger's Law" requiring that all bloggers with a daily audience of more than 3000 register on a national list and follow media regulations for fact-checking their posts. This package also included an "anti-encryption" law, requiring that all encrypted services provide government backdoors, and a "user data storage" requirement, that went into force in 2016, requiring that all Internet sites and platforms that collect data from Russian users must store that data for three years and provide for government access.

Such laws, though almost never systematically enforced, create significant chilling effects both for content producers and intermediaries as well as providing legal grounds for subsequent blockings or prosecutions.

Online media outlets and social media platforms face the threat of potential financial takeovers and pressures to swap editors, CEOs, or other key personnel, if they fail to bow to content restriction pressures. Pavel Durov, the founder of Russia's most popular social network, VKontakte, left the country in April 2014 after being fired as CEO and forced to sell his shares in the company leaving it majority owned by oligarchs close to the Kremlin. Durov publicly stated that the conflict had resulted from his unwillingness to disclose user information or block pages relating to Alexei Navalny's anti-corruption campaign and the conflict in Ukraine. Durov subsequently founded the messaging app company Telegram, which was officially blocked (though not at all successfully) prompting protests in 2018 after refusing to turn over encryption keys.

Changes in surveillance laws and capabilities have been an important area of increased government control in the post-2012 period – though it is not always clear to exactly what extent and ends the collected data is being utilized. Internet service providers (ISPs) and social media platforms alike have faced pressure to quickly implement new requirements such as the purchase and installation of surveillance equipment on their networks or the storage of and government access to all user metadata and communications. Russia's mass surveillance system, SORM, is grounded on a legal framework allowing for the "lawful interception" of communications by a number of KGB-successor security organs and other government bodies. It also involves particular technological systems and infrastructures used to implement the data storage and access. Both the SORM regulation and technology have received recent enhancements. Whereas earlier SORM-2 systems had only operated at the ISP level, an August 2014 decree required all social media platforms operating in Russia to install SORM monitoring equipment. The new SORM-3 system, announced also in 2014, was to permit the storage of all communications and tracking of data streams by particular users and IP addresses.

In the Russian approach to information control, in addition to surveillance and legal and extra-legal pressures, new forms of pro-regime content mass-production and narrative manipulation as well as the limited use of plausibly deniable cyberattacks and hacking play critical roles in efforts to undermine and marginalize the voices of opposition movements and leaders, while also shaping broader public opinion without a sense of dramatic restriction. The leveraging of youth organizations (such as Nashi), third-party botnets, independent hackers, contracted video-producers, and pro-regime bloggers in coordinated actions provides a further degree of deniability of government involvement. Bots, trolls, leaks of compromising or manipulated content, DDoS attacks causing temporary "technical failures," and other difficult-to-attribute techniques are combined with occasional legal prosecutions or site-blockages for exemplary offenders under vague laws and mass digital surveillance, creating an overall online environment which still appears relatively unrestricted – with the ability to produce and access wide varieties of content, including content critical of the government – but in which the government exerts significantly more control over the overall development of content and narratives.

In the realm of content production, Russia has shown significant experimentation in its effort to gain greater control over domestic opinion and dampen sources of political instability. While originally seeking to sway public opinion primarily through television content, the approach has been updated in recent years to adjust for the growing domestic political significance of Internet content consumption. The new 2016 version of the Russian "Doctrine of Information Security" explicitly discussed the roles of the Internet and social media as well as other mediums for information production and consumption. While some forms of propaganda and tools of narrative manipulation are repeated across all platforms in coordinated efforts, other techniques appear to have been developed explicitly to take advantage of the affordances and vulnerabilities of the digital media ecosystem.

Russian content production and manipulation efforts often pay careful attention to framing and agenda-setting, playing off of existing biases, identities, societally-resonant symbols, and the manipulation of emotion.[3] In some cases, efforts aim to promote particular narratives. In others they plant numerous alternatives to existing narratives sowing confusion and uncertainty (e.g. "who downed MH17?" "who was behind chemical weapons attacks in Syria?"). They also interrupt and distract politically critical conversations, dilute potentially critical discourse contexts with fun apolitical content[4] (e.g. the discussion surrounding politically salient hashtags), or seed different content into different echo chambers to further exacerbate existing tensions. These techniques can aim to drown out criticism or break potential protest coalitions, preventing critical discourse from leading to political mobilization without the need for frequent censorship.

Relationship to International Information Conflict

[3] In some cases, framing and agenda setting appear to be given particularly systematic attention, even utilizing the broader global information flows to the regime's benefit. In the period following the White Ribbon movement's mass mobilization of a diverse coalition to protest regime corruption and electoral fraud, the imprisonment of members of the feminist punk girl band Pussy Riot (for staging a protest inside a cathedral), and the ratcheting up of pressure on LGBT groups both seemed calculated to draw attention to the less traditional values expressed by small subsets of the protest movement. This attention – reflected back and magnified through Western civil society and governmental attention and outrage – helped to reframe the protest movement as one concerned primarily with these progressive issues, weakening the cross-coalition bonds between different protest participant groups and reducing the resonance for the majority of participants who had mobilized around economic and political rights. At the same time, moderate protest mobilizations in Moscow concerned with peace with Ukraine and media freedom received little such coordinated media attention.

[4] This sometimes includes content which could pass as either satire or propaganda and thus is spread by supporters and opponents. Such content sometimes plays to pop-cultural tropes, memes, and content-production patterns, blending with other popular content and even inspiring copycat content production.

The Russian approach to domestic control of information within society has direct applicability to the leveraging of information operations in international political and military competition. It also is closely tied conceptually. Throughout the 2000s, as concern about the existential threat to regime survival posed by mass protest events grew, the leadership increasingly came to worry about the roles of transnational information flows as part of military and strategic competition and as potential sources of domestic political instability. While the U.S. and democratic allies promoted "Internet freedom" as a distinct issue from growing attention to national cybersecurity and the military cyber domain, the Russian understanding of "information security" and international information aggression subsumed both the transnational networked flows of media and information and the networked computer systems and data that were generally the focus of cybersecurity analysis.

This consideration is clear in an often-quoted article by Russia's then chief of the general staff, General Valery Gerasimov, that focused on Arab Spring type events as part of an analysis of the current military-technological and geopolitical threat landscape, and suggested that "broad use of political, economic, informational, humanitarian, and other non-military measures – applied in coordination with the protest potential of the population" were playing increasingly significant roles in contemporary forms of strategic international competition. Speaking of threats posed to Russia, he stressed Russia's need to also utilize such combined efforts, engaging in "cognitive-psychological" and "digital-technological" forms of influence. Suggesting that strategic goals could be achieved with little resort to armed conflict[5] through influencing perceptions and decision-making processes, he stressed the importance of "information spaces" and the possibility of exploiting asymmetric vulnerabilities, even against more militarily-powerful adversaries.

Evidence today suggests that Russia utilizes information operations abroad, both in regional and international theaters, at levels targeting individuals, groups or entire populations. These are applied to undermine credibility or intimidate, plant particular narratives and distract from others, sow confusion and uncertainty, exacerbate divisions, galvanize protest, and slow or influence decision-making processes. Goals appear to include efforts to influence elections, undermine support for political parties and candidates, support extremism and polarization, and undermine the legitimacy of institutions not aligned with Russian foreign policy. Techniques sometimes involve both technical (hacking, malware) and informational (content) components, including actions such as leaks of compromising material, DDoS attacks, and website defacements. They often take advantage of plausible (or even implausible) deniability, and can occur during peacetime, grey zone (sub-threshold) conflicts, or wartime, in combination with special operations, direct military action, or diplomatic interaction. In addition to state-organ-led efforts, Russia appears to also sometimes leverage hacktivists, youth organizations, criminal networks, and paid troll farms (e.g. the Saint Petersburg-based Internet Research Agency) as state proxies to conduct operations, aiming to obfuscate direct governmental links.

As elements of international geopolitical competition, these techniques draw on a long tradition within Russian and Soviet military strategy. Soviet "active measures" and the concepts of "maskirovka" and "reflexive control" in Russian military theory each involve the use of information and deception, ambiguity and illusion, and deniable and indirect activities, for the purposes of psychological manipulation and asymmetric influence. The more recent cyber-enabled information operations are differentiated, however, by the ubiquity and affordances of the digital technologies being utilized. Using the new technologies, significant influence effects can be achieved remotely, quickly, on scale, and at relatively low cost – at least in theory. Some of these cyber-enabled information and influence operations are undoubtedly more

[5] He suggested a 4-to-1 ratio of non-military to military operations.

effective than others. As in the domestic sphere, there is evidence of experimentation to develop more effective uses of the current tools for information manipulation.[6]

The international applicability of aspects of the Russian domestic model for information control suggests that ongoing learning and experimentation within authoritarian regimes will have continuing relevance to international information contestation. This also brings into question the Cold War era assumption that democracies, having less to fear from public discourse and free expression, are always more resilient to international flows of information than are non-democratic regimes. Democratic countries may in fact have some important vulnerabilities that are different and greater in the face of the new information operation techniques.

So What? And What Next?

The Russian model of Internet control should not be reified. It has emerged out of ongoing experimentation, and sometimes seems as much shaped by opposing internal inclinations or by a failure to adequately implement more robust censorship models as by an intentional effort to maintain some semblance of democratic legitimacy. Why then, if at all, is the distinctiveness of the Russian approach worth noting? There are at least two significant reasons.

The first is the applicability of some subset of this model's features to regional and international theaters. This means that experimentation and learning around information control at home can drive advances in "political" or "information" warfare capabilities in international competition. The second is the potential broader diffusion of this model – both the domestic and international elements – to countries for which a sophisticated censorship approach might, for various reasons, not be within grasp.[7] The continued success and diffusion of the model's domestic approach promises a potential path forward for hybrid regimes in the digital age. The demonstration of its utility in regional and international conflict is likely also to serve as inspiration for many copycats.

But this leaves several unresolved questions. Can this model continue? Is it possible, long-term, to retain as much (or sufficient) control over public opinion through content production, surveillance, and limited censorship as through ubiquitous censorship? The continuing success of this approach will require ongoing innovation. So the answer might depend on the next steps. How is this likely to develop further in the near future? Should we expect an eventual convergence with or continued distinction from the Chinese model?

---

[6] While early examples occurred in the 2000s, including operations during diplomatic and military conflicts with Estonia (2007) and Georgia (2008) respectively, more recent events, particularly since the 2014 beginning of the conflict with Ukraine and surrounding elections in the United States and Europe, show ongoing experimentation and learning. The more recent campaigns have used the hacking and leaking of confidential information to manipulate media discussion (e.g. DNC hack), leveraged major social media platforms through bots and other fake accounts to disseminate content, and used micro-targeting and advertising technologies and the manipulation of existing fringe or activist echo chambers and group sites to introduce false information or alternative narratives, exacerbate social divisions, influence public discourse, and catalyze real-world protest events.

[7] The domestic model is likely of particular ease to emulate across the former Soviet region, as these states share legal and institutional legacies, participate in common regional organizations, and also often share overlapping media markets and Internet resources. There is already evidence of significant diffusion of aspects of this model across the region, including, for example, the SORM surveillance infrastructure and accompanying "lawful intercept" legal frameworks permitting access for KGB-successor organs, national security frameworks focused on the role of information, and many similar hacking and content production and manipulation tactics. The model is also of clear merit to other hybrid regimes, which wrestle with the same conflicting pressures, however. And some aspects of this approach are likely to prove valuable to states of various non-democratic regime types that for technical, financial, human capital, or organizational reasons have more adequate capacities to implement an approach that relies less on technical systematic censorship and more on the prosecution or censorship of exemplars, use of broad legal rules, and content production.

Of particular significance for the future of digital authoritarian models and global information-power competition will be the interrelated roles of artificial intelligence (AI) and big data. Advances in machine learning are now driving breakthroughs in a variety of technologies relevant to online discourse and its monitoring, censorship, or emulation. As demonstrated by Cambridge Analytica, AI and big data permit ever-more-precise forms of micro-targeting – whether for advertising or propaganda. Algorithms also now permit the production of increasingly inexpensive and realistic deep fakes – fabricated lifelike audio and video files that can make it appear that someone said or did something that they did not. Improvements in sentiment analysis and natural language processing allow better analysis of emotion – useful for targeting and engaging individuals and populations. Meanwhile, chatbots are finally passing the "Turing Test," with some experiments showing subjects unable to differentiate between interactions with real people and computer agents in certain settings. In states which have struggled to implement systematic Internet and information controls, shouldn't these tools permit more ubiquitous censorship and more perfect law enforcement?

In a September 2017 speech, Vladimir Putin noted the importance of AI. "Artificial intelligence is the future," he told the nation's students, "not only for Russia, but for all humankind." A great deal of attention has focused on the Chinese government's access to large quantities of data – critical to the training and effective use of AI algorithms. But with all the input from the SORM surveillance systems and recent data storage requirements, the Russian government is likely also to have significant data with which to experiment.

How might the role of AI and big data in information control look different in a Russian context? One could imagine this the solution to all of Russia's censorship and enforcement woes. But if the content production and limited censorship approach continues to prove effective, it seems more likely that Russia would use AI in ways consistent with that model: more precise micro-targeting, more emotional manipulation, more believable and impactful propagandistic content – and, importantly, the use of these same tools at home and abroad. This is something worth anticipating and preparing for…