DISINFORMATION AND DIGITAL DEMOCRACIES IN THE 21ST CENTURY

NATO ASSOCIATION OF CANADA
FALL 2019

The NATO Association of Canada
Disinformation and Digital Democracies in the 21st Century

Senior Content Editor:
Dr. Joseph McQuade

Assistant Editors:
Tiffany Kwok
James Cho

Design:
James Cho

Contributing Authors:

Ryan Atkinson
Sebastian Bay
Josh Campbell
Elisha Corbett
Abigail Curlew
Victoria Heath
Ian Litschko
Jazlyn Melnychuk
Dr. Jeffrey Monaghan
Christian Picard
Dr. Regina Rini
Janis Sarts
Guna Šnore
Dr. Heidi Tworek

The NATO Association of Canada—Association Canadienne pour L'OTAN
Chairman of the Board: Hon. Hugh Segal, OC, OOnt
President: Robert Baines, CD, MA

The NATO Association of Canada is an independent, non-profit, non-governmental organization dedicated to the idea that the transatlantic relationship between Canada, the United States, and the nations of Europe is of critical military, economic, and cultural importance to Canadians. The Association's mandate is to promote a broader and deeper understanding of international peace and security issues relating to NATO.

48 Yonge St, 610
Toronto, ON, Canada, M5E 1G6
Phone: (416) 979-1875
Facsimile: (416) 979-0825
Email: info@natoassociation.ca

FOREWORD

In recent years, disinformation (known colloquially as "fake news") has come to dominate the public consciousness in a way that has not been seen since the end of the Cold War. Digital technologies have penetrated our societies and personal lives to such an extent that it is easy to forget that the world of smartphones is less than two decades old. Still, as the past few years have shown with devastating clarity, the dangers of online data breaches, disinformation, and mass surveillance have very real consequences in the offline world.

Having just concluded a federal election in which the Communications Security Establishment (CSE) warned of the potential for unprecedented interference, Canada is by no means immune to the risks posed by disinformation spread through new digital platforms. Now that the dust has settled and the election has concluded, it is useful to take stock of the relationship between disinformation and democracy, in both domestic and global contexts. While most of the articles in this collection focus primarily on the implications of disinformation and digital insecurity in Canada and in other NATO member-states and allies, we hope that this issue can provide insights that are useful in a range of global contexts. In a world increasingly connected by social media, digital communities, and smartphones, the impact of disinformation and other threats to digital democracies cannot be understood other than in transnational terms.

Even just a handful of examples can highlight the global scope of the problem, and the myriad forms it can take. In 2015, military officials and Buddhist nationalists in Myanmar used fake stories and videos on Facebook to incite pogroms against the Rohingya ethnic minority, displacing around a million civilians and creating a humanitarian catastrophe that persists today. In 2016, Russian trolls bombarded the internet with divisive stories and rhetoric that helped polarize the US presidential election and the UK Brexit referendum, contributing to populist surges in both countries. Between May 2015 and December 2018, 44 people from minority and low-caste backgrounds were killed and another 280 injured across India as a result of "cow vigilantes" who seek out and lynch those accused of slaughtering cows, often as a result of false rumours spread through digital platforms such as What's App. In Iraq and Syria, the Islamic State has famously deployed social media and digital networks as a means of recruiting "foreign fighters" and promoting attacks by "lone wolf" extremists living around the world. Radicalized by a shadowy online "incel" culture, young mysoginists have committed devastating acts of mass murder in cities ranging from Isla Vista to Toronto. At the same time, online conspiracies promoting a supposed connection between vaccines and autism have caused vaccination rates to plummet in developed countries, leading to a resurgence of preventable illnesses such as measles.

All the while, tech giants such as Google, Facebook, Microsoft, Apple, and Amazon hoover up unimaginable quantities of user data, which is then sold on "behavioural futures markets" to ensure that those wishing to sell us products – whether of the consumer or the political variety – can do so through the use of predictive algorithms that often know us better than we know ourselves. While the global surveillance capabilities of America's National Security Agency (NSA) were made famous by former intelligence contractor Edward Snowden back in 2013, the public's awareness of data privacy remains limited to this day. Unwieldy terms of service agreement pages and ubiquitous cookies that track web traffic on virtually all major sites make it almost impossible for even the most engaged citizens to truly understand who does or does not have access to their personal information. Meanwhile, the massive quantities of data available to the Chinese government – generated both by China's own sizeable population and by those of Eurasian and African countries swiftly developing

their digital infrastructure thanks to the trillion-dollar investment of the Belt and Road Initiative – are facilitating new breakthroughs in machine learning, facial recognition, and Artificial Intelligence that could soon make older surveillance systems seem positively archaic by comparison.

For all of these reasons and more, issues of disinformation and digital citizenship have become inseparable from questions of democracy in the 21st century. Just as urbanization and the rise of print media in the nineteenth century profoundly shaped the rise of democracy in the modern world, the future of democracy will almost certainly be defined largely by the processes currently unfolding on computer screens across the globe. Still, we should be careful in drawing too neat a line between the eras of the analog and the digital. In most cases, the technological developments of our age have reshaped existing social patterns and hierarchies of power, rather than creating wholly new ones from scratch. Despite the early promise of the internet as a democratizing force that would give everyone equal access to information and influence, marginalized and disadvantaged communities remain disproportionately disenfranchised in the digital realm.

The relevance for NATO and its allies is inescapable. The most successful multilateral military alliance in modern history, NATO is comprised of a constellation of democracies spanning both sides of the Atlantic, with key allies across the Pacific, South Asia, and the Middle East. In conventional terms, information warfare, cyber terrorism, and hacking have obvious implications for the security of NATO members, a fact reflected in Secretary General Jens Stoltenberg's confirmation that the famous mutual defence clause of Article 5 could be triggered by a cyber-attack. Aside from the instances of electoral interference cited above, Russia under Vladimir Putin has launched substantial information warfare operations along its Eurasian frontier, from Ukraine to Syria. To counter these operations, NATO and its allies have launched a range of new initiatives, including the G7 Rapid Response Mechanism and the Riga-based NATO Stratcom Centre for Excellence.

These initiatives make no secret of the fact that this policy-oriented research must be directly linked to grassroots organizations, academics, and civil society groups who share the aim of strengthening public resiliency and building a firm foundation of digital literacy. While it is true that many of the most polarizing digital disinformation campaigns in recent years bear the trace of Russian fingerprints, these campaigns are only successful when they tap into existing fault-lines produced by decades, and sometimes centuries, of social unrest or political animosity. Grappling with the immense societal and security challenges posed by the technological transformation currently underway will only be possible when we understand that the vulnerabilities, dangers, and dispossessions of the digital world have built on a foundation with deep historical roots.

The articles in this issue address the themes of disinformation and digital democracy from a range of perspectives. Bringing together new and original findings from established experts in the field and some of the country's most promising early career researchers, from both inside and outside of academia, the issue seeks to simultaneously expand and refine our understanding of how digital technologies have paved the way for new possibilities of emancipation and oppression, information and disinformation, authoritarianism and democracy.

Dr. Joseph McQuade
Editor-in-Chief, NATO Association of Canada
RCL Postdoctoral Fellow, Asian Institute,
Munk School of Global Affairs and Public Policy, University of Toronto

TABLE OF CONTENTS

Threats to Democracy: Technological Trends and Disinformation

Janis Sarts, *Director, NATO Strategic Communications Centre of Excellence*
Sebastian Bay, *Chief Analyst*
Guna Šnore, *Senior Expert*
Jazlyn Melnychuk, *Intern*

What is the StratCom COE?

The NATO Strategic Communications Centre of Excellence is a multi-nationally constituted and NATO-accredited international military organisation. It is not part of the NATO Command Structure, nor is it subordinate to any other NATO entity. The NATO StratCom COE, based in Riga, Latvia, contributes to improved strategic communications capabilities within the Alliance, Allied nations, and Partner Nations.

The Importance of Being Proactive

Although the Canadian federal elections will be complete by the time of this publication, the conversation regarding protecting the integrity of the electoral process must continue. As our allies and partners around the world enter election seasons, studying each case and drawing lessons learned can be the difference between a strong defence and a democratic backslide. Democracies have long been accused of being reactive rather than preventive in the digital age. Now, our democracies and their processes are becoming digital, meaning the threats posed by hostile actors must be prioritized in order to safeguard our values.

The risk of encountering malicious foreign influence during the elections can be reduced with a proactive approach that is active long before elections begin. Canada has implemented several important measures to address electoral interference. Bill C-76, passed at the end of 2018, prohibits advocacy groups from utilizing foreign funding and seeks to create transparency around the purchase of political ads on social media.[1] However, as tactics and technology evolve, so too must policy, making the consistent tracking of trends essential. This article will discuss examples of recent foreign influence campaigns, identify various types of electoral inference, and explore future technological trends in disinformation.

Examples of Recent Influence

A recent study from Princeton University identified 53 targeted Foreign Influence Efforts (FIEs) between 2013 and 2018, assigning responsibility for 72% of these campaigns to Russia, with 29 operations in 2017 alone.[2] Iran is also becoming increasingly active, targeting the United States and United Kingdom, among others. In 2017, Russia-linked hackers attacked government ministries and an anti-Russian political party in Norway. In the Netherlands in 2017, domestic intelligence officials reported that foreign

[1] Panetta, Alexander, and Mark Scott. "Unlike U.S., Canada Plans Coordinated Attack on Foreign Election Interference." POLITICO, September 4, 2019. https://www.politico.com/story/2019/09/04/canada-foreign-election-meddling-1698209.

[2] Martin, Diego A, and Jacob N Shapiro. "Trends in Online Foreign Influence Efforts." Dissertation, Woodrow Wilson School of International and Public Affairs, 2019.

countries, notably Russia, tried hundreds of times to penetrate the computers of government agencies. During the European Union election cycle of 2019, Microsoft reported 104 breach attempts to government institutions and non-profit organizations, with a focus on gaining access to employee credentials and delivering malware.[3]

While Canada has fared well compared to other allies, it has faced its own struggles with disinformation. In 2018, Russian troll accounts on Twitter with a right-wing lean mentioned Canadian topics including the Keystone XL pipeline, asylum seekers, and the Quebec City Mosque shooting nearly 8,000 times.[4] The intent of these attacks was to polarize public opinion. Most recently, an October 2019 poll suggests that Canadians are being influenced by disinformation regarding immigration. The majority of respondents believed 64% of immigrants were from Africa and the Middle East, when in reality this region accounts for 12% of immigrants.[5]

An analytical framework for understanding electoral interference

Election interference aims to influence the outcome of an election, to undermine trust in the election, or to use the election to achieve other goals, such as undermining democracy, internal cohesion or influencing how a country is perceived externally.

These effects can be accomplished by influencing a) the election as an administrative process, b) the will and ability of voters to participate in the election, and c) the election as a political process. These three components should be seen as interconnected processes that can be influenced via various activities, ranging from targeting the election infrastructure to manipulating the political debate.

Threats against the election as an administrative process

Without the public's trust that elections can deliver a credible result, a cornerstone of democracy is at risk. Even without antagonists seeking to influence them, elections are complex undertakings with numerous risks linked to the legal, operational, technical, political, and security aspects of electoral processes.

Threats against the will and ability of voters to participate in elections

In order to conduct legitimate, free and fair elections, it is necessary to protect the will and ability of the population to participate in them. This entails ensuring equal access to correct information about where, when and how citizens can vote. In recent elections, there have been attempts to spread false information about how they are carried out. Antagonists have tried to get voters to not participate by spreading false information about where, when and how citizens can vote. We have also seen attempts at voter intimidation, with the aim of undermining voter participation.

Threats against the election as a political process

Illegitimate information influencing activities are different from legitimate communication activities in that they are *deceptive:* they involve falsehoods in some way or other, they have the *intention* to exploit vulnerabilities to benefit a foreign power or its proxies, they seek to *disrupt* constructive debate, and they *interfere* in debates or issues in which foreign actors have no legitimate role to play.

[3] Mr. "Figure of the Week: 104." EU vs DISINFORMATION, April 3, 2019. https://euvsdisinfo.eu/figure-of-the-week-104/.

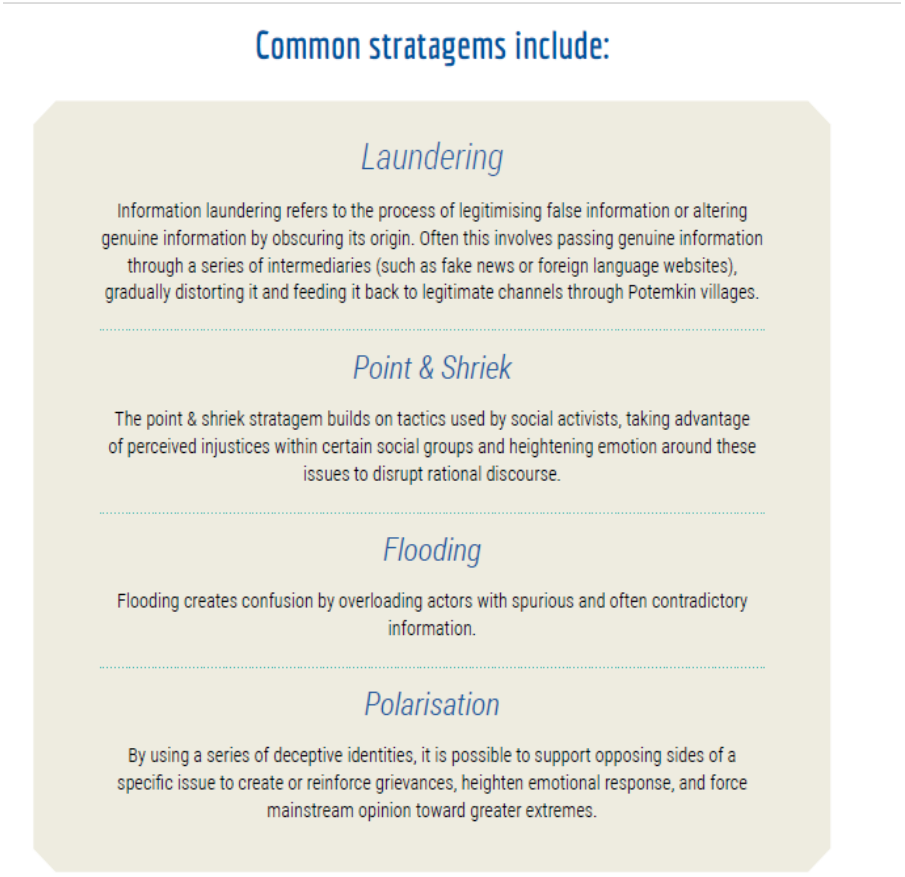
[4] Martin, Diego A, and Jacob N Shapiro. "Trends in Online Foreign Influence Efforts."

[5] Rogers, Kaleigh. "Canadians' Misperceptions about Immigration Reflect Disinformation Online: Experts." CBCnews. CBC/Radio Canada, October 13, 2019. https://www.cbc.ca/news/politics/immigration-disinformation-election-2019-poll-1.5316934.

Hostile actors have previously tried to interfere with the political process by means of cyberattacks directed at political parties, the publication of stolen and manipulated information, targeted micro ads against a vulnerable target audience, paid and automated manipulation through social media, and so on. Attempts to interfere in the political process using subversion, proxies and other forms of illegitimate means to distort a political process falls within this category of interference.

Stratagems

Individual methods and techniques for election interference are rarely used in isolation. Rather, influence operations and campaigns most often combine a multitude of methods and techniques into complex chain of events, or stratagems. While such combinations are theoretically infinite, some stratagems are frequently encountered in contemporary influence operations.



Technological Trends and Future Manipulation

As fast as you can implement new policy – hostile actors are working to get around them. The true ramifications of the 2016 US Election interference are still being realized, with a high degree of uncertainty felt throughout the Alliance. It is likely that despite new steps among allies and partners, the methods

supporting FIEs will only evolve to circumvent these measures. The aforementioned stratagems will not only evolve to evade policy, but change with technological shifts. Three key threats facilitated by new technological trends include enhanced deep fake technology, encrypted communications platforms, and enhanced accessibility and commercialization of disinformation campaigns.

A concept that is eliciting growing attention is deepfakes, whose sophistication and ease of access continues to grow. In May 2019, Samsung's Moscow Artificial Intelligence Lab created a Generative Adversarial Network (GAN, i.e. a set of deep learning algorithms which facilitate deepfakes) whose model required only one image of an individual to input their likeness into a video, with previous models requiring hundreds of thousands of images.[6] Another emerging challenge comes from new AI-driven technology capable of writing political speeches and articles, now being termed "deeptext."[7] Deeptext allows AI to draft full texts with human-like complexity such as imitating UN speeches on sensitive topics such as disarmament and refugees.[8] In response, Harvard researchers created an AI tool that spots fake AI-written text with a success rate of 72% which can be used by the general public.[9]

Another area of growing concern is the increased use of closed-network communications networks to avoid growing protection measures. The key platforms in this domain are WhatsApp and Telegram, where it is extremely difficult to monitor the dissemination of disinformation, limiting the work of fact-checkers and researchers. As messages are end-to-end encrypted, locating the source of disinformation once it is being forwarded is nearly impossible. In January 2019, WhatsApp reduced the number of groups content could be forward to from 20 groups to 5 groups, coming down significantly from the original limit of 256.[10] However, a recent study found that 20% of disinformation through WhatsApp was still going viral and reaching its full audience, with total forwarding decreasing by only 25%.[11] Another danger of encrypted networks is that the disinformation is often propagated through intimate networks from people we know and trust, likely making the impact on opinion/behavior more significant.[12] Disinformation through WhatsApp has been used to manipulate voters in Nigerian and Brazilian elections, and has even provoked fatal attacks in India.[13]

While disinformation was first perceived as a complex, coordinated act by foreign governments, the technology and tactics used to implement influence campaigns have

[6] Littell, Joe. "Don't Believe Your Eyes (or Ears): The Weaponization of Artificial Intelligence, Machine Learning, and Deepfakes." War on the Rocks. Texas National Security Review, October 7, 2019. https://warontherocks.com/2019/10/dont-believe-your-eyes-or-ears-the-weaponization-of-artificial-intelligence-machine-learning-and-deepfakes/.

[7] Hundeyin, David. "This New 'Fake Text' AI Is Even More Terrifying Than Deepfakes." CCN.com. CCN Markets, July 7, 2019. https://www.ccn.com/fake-text-ai-worse-deepfakes/.

[8] Hao, Karen. "You Can Train an AI to Fake UN Speeches in Just 13 Hours." MIT Technology Review. MIT, June 8, 2019. https://www.technologyreview.com/f/613645/ai-fake-news-deepfakes-misinformation-united-nations/.

[9] Knight, Will. "A New Tool Uses AI to Spot Text Written by AI." MIT Technology Review. MIT, July 26, 2019. https://www.technologyreview.com/f/614021/a-new-tool-uses-ai-to-spot-text-written-by-ai/.

[10] Chen, Angela. "Limiting Message Forwarding on WhatsApp Helped Slow Disinformation." MIT Technology Review. MIT, September 30, 2019. https://www.technologyreview.com/f/614435/whatsapp-disinformation-message-forwarding-politics-technology-brazil-india-election/.

[11] ibid

[12] "How WhatsApp Is Used and Misused in Africa." The Economist. The Economist Newspaper, July 18, 2019. https://www.economist.com/middle-east-and-africa/2019/07/18/how-whatsapp-is-used-and-misused-in-africa.

[13] ibid

become widely accessible. A study conducted by *EU vs Disinfo* found that the cost of contracting a disinformation post costs as little as $8 and can be purchased by anyone. They also found that "[b]ased on the importance of the outlet, the black market prices for publishing a disinformation article ranged from $180 at business websites and rose up to $8,360 for Reuters, $13,370 for Mashable and $49,440 for Financial Times," showing the sophistication of available services.[14] The commercialization of disinformation campaigns incorporates a new profit angle combined with the political motivations of the client, creating a perfect storm of incentives for malicious actors of all types. Our forthcoming study purchased fake interactions across four platforms and reported them as inauthentic to view the response of social media companies. The study revealed that these services are quite effective, with 95% of fake accounts still active three weeks after they were reported.[15] A lack of transparency in political advertisements and disinformation attacks against candidates pose severe risks to the independence of the democratic process, which become more likely as the black market for disinformation grows.

Demand More –from Everyone

For social media companies, experts at the COE have recommended that companies increase monitoring of impersonations of government accounts, mandate ad transparency, and develop better algorithms for recognizing non-organic manipulation of user content aimed at affecting the perceived popularity of a particular view.[16] We must be vigilant to the ways in which malicious actors will seek to circumvent these policies, and be ready to adapt with technological advancements. We must learn how to evolve our electoral regulations alongside the changes in the information environment, or risk leaving our democracies vulnerable.

While governments and companies bear significant responsibility, we must not forget the importance of vigilance at the individual level. Even with fact-checkers and social media companies responding to disinformation, debunking lies can be extremely time consuming. This is problematic for election periods, where a strategically-timed deepfake can alter voter behavior. Fundamentally, one of the best defences against electoral interference is a critical, skeptical, and informed voter, which is why increasing digital literacy and awareness should be a government priority.

Study after study has shown that current efforts are largely ineffective, with innovation outpacing law, but there is no easy fix to a problem that crosses political, societal, economic, and even military lines. Effectively addressing disinformation in the electoral context requires a whole of society effort that holds to account social media companies and political actors alike. Ultimately, civil society actors, voters, and politicians need to demand more –of companies, of governments, of each other, and even of themselves.

[14] "Figure of the Week: 8." EU vs DISINFORMATION, October 8, 2019. https://euvsdisinfo.eu/figure-of-the-week-8/.

[15] Forthcoming StratCom COE Research.

[16] Bay, Sebastian, and Guna Snore. "Protecting Elections: A Strategic Communications Approach." *NATO Strategic Communications Centre of Excellence*, June 2019, 1–22. https://www.stratcomcoe.org/protecting-elections-strategic-communications-approach.

Social Media Disinformation and the Security Threat to Democratic Legitimacy

Dr. Regina Rini, C*anada Research Chair in Philosophy of Moral and Social Cognition, York University*

Everyone now knows about Project Lakhta, the Russian political interference operation conducted largely through fake social media profiles and aggressive disinformation. In February 2018 the US Justice Department indicted 13 affiliates of the St. Petersburg Internet Research Agency (IRA), alleging that they had conspired to "spread distrust toward the candidates and the political system in general".[17] Other reports suggest similar operations were conducted against the United Kingdom and Germany. International news headlines put democratic citizens on notice: their politics are being manipulated by foreign bots and trolls.

But most discussion of social media interference has been misleading in at least two ways. First, commentators tend to focus on the tricky empirical question of whether Russia succeeded in changing voting outcomes – did it throw the 2016 US election to Donald Trump, or tip the Brexit referendum toward Leave? This misses the crucial point that these operations seem to be aimed less at affecting outcomes and more at stoking heightened domestic discord and partisan distrust. Second, many analyses treat social media interference as simply another species of political disinformation, akin to traditional propaganda. But this misses the most dangerous feature of social media interference operations. When citizens unwittingly spread disinformation by sharing social media posts from fake accounts they become *complicit* in their own deception, undermining their standing to credibly govern one another. In other words, these operations are intended to weaken the legitimacy of democratic government itself. And, unlike other forms of disinformation, their effects are most severe *after* they have been exposed, as citizens become aware that their fellow democratic deliberators are prone to being duped by foreign manipulation.

I will expand on this point in three parts. First, I will highlight the ways that democratic legitimacy depends on citizens' individual credibility. Then I will show how this credibility is systematically undermined by social media interference, especially when it has been exposed. I will conclude with policy recommendations for NATO member states.

The epistemic preconditions of democratic legitimacy

In recent decades, political philosophers have increasingly noted the *epistemic* features of democratic legitimacy. Since laws intrude on individuals' freedom, the perceived legitimacy of state authority depends on its competence. A state claims legitimacy by representing itself as best equipped to determine for citizens their political obligations.[18] Democratic legitimacy therefore depends on the assumption that citizens themselves, acting through public debate and selection of representatives, are collectively better at figuring out their obligations than they would be on their own. In particular, it is assumed that citizens possess

[17] *United States of America v. Internet Research Agency* federal indictment. February 2018. 6. Text available at https://www.justice.gov/opa/press-release/file/1035562/download

[18] See Joseph Raz (1986), *The Morality of Freedom*. Oxford: Oxford University Press.

adequate *epistemic competency* – that they are sensible in how they acquire, share, and debate knowledge – and that public debate and democratic voting enhance rather than detract from collective reasoning.

Overtly epistemic questions are visible in many issues debated by democratic publics. For example: what role should expert opinion play in public deliberation? Should exclusive epistemic frameworks (e.g. faith assumptions) have full evidential status? How should policy account for uncertain scientific knowledge (e.g. details of the impact of climate change?) Should journalists be granted special immunities or protections to facilitate public information gathering?

Political theorists actively dispute the extent to which democratic systems *succeed* in fulfilling their epistemic functions. Philosopher David Estlund argues that fair democratic procedures provide citizens reason to comply even with laws they personally do not regard as wise. Political theorist Hélène Landemore goes further, arguing that (some) existing democratic systems do *in fact* improve public reasoning. However, other theorists dispute these claims. Philosopher Alex Guerrero argues that, since real-world politicking often directs attention to irrelevant electoral drama, better decisions would come from a 'lottocratic' system in which randomly-drawn assemblies of citizens made laws. Jason Brennan goes farther still, arguing that many citizens do not in fact possess sufficient epistemic competence to make their authority over others legitimate. Instead, Brennan claims, we should implement 'epistocracy', in which citizens who perform well on tests of factual knowledge wield more votes than others.[19]

The radicalness of these latter proposals shows the extent to which we currently assume democratic citizens deserve epistemic authority. This assumption means that democratic states are vulnerable in a way that authoritarian states are not. The legitimacy of an authoritarian state (to whatever extent there is such a thing) does not depend on the epistemic competence of individual citizens, since the authoritarian state does not typically consult its citizens' judgment. This means that attacks on citizens' epistemic competence affect democratic states much more severely.

This contrast was noted by the early 20th century Russian political theorist Ivan Ilyin, who regarded democracy as a fundamentally corrupt form of government, prone to empowering liars and fools. Ilyin advocated for minority factions to expose democracy's foolishness to its own citizens, eroding legitimacy until the democratic system failed and could be replaced by totalitarianism. Ilyin's objectives were directed primarily at domestic politics, but historian Timothy Snyder has documented Ilyin's influence on contemporary Russia's international strategy, including its social media interference operations.[20] The vulnerability of democratic legitimacy can be weaponized by authoritarian states prepared to amplify democratic citizens' suspicions of their own compatriots.

Social media disinformation makes democratic citizens complicit in their own befuddlement

In this section I will show how recent Russian social media interference operations target the legitimacy of democratic states. The scope of these operations is striking in itself; in late 2017

[19] David Estlund (2008), *Democratic Authority: A Philosophical Framework*. Princeton: Princeton University press; Hélène Landemore (2012), *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many*, Princeton: Princeton University Press; Alexander A. Guerrero (2014), 'Against Elections: The Lottocratic Alternative', *Philosophy and Public Affairs* 42(2):135-178; Jason Brennan (2017), *Against Democracy*, Princeton: Princeton University Press.

[20] Timothy Snyder (2018). *The Road to Unfreedom: Russia, Europe, America.* New York: Tim Duggan Books.

Facebook acknowledged that as many as 126 million Americans were exposed to 2016 electoral content posted by IRA operatives masquerading as fellow citizens. Twitter identified (and deleted) 2,752 fake IRA accounts. Researchers at the University of Edinburgh identified 419 IRA Twitter accounts participating in debate over the Brexit referendum. These techniques continued beyond the 2016 votes and seem to have been added to the arsenals of other authoritarian states; in August 2018 Facebook announced that it had deleted an additional 652 fake accounts linked to Russia or Iran.[21]

These operations undermine the epistemic presuppositions of democratic legitimacy in at least two ways. First, they trick citizens into actively spreading disinformation (through Facebook shares and Twitter retweets), making them personally complicit in their own befuddlement. Sharing of 'fake news' erodes norms of epistemic accountability among citizens and amplifies the divisive effects of partisanship.[22]

When citizens observe one another's complicity, they acquire evidence against the assumption that their co-citizens possess epistemic competence, weakening the perceived legitimacy of democratic decision-making.

Second, the use of fake *accounts* – invented or stolen identities purporting to be citizens of democratic states – threatens to make democratic citizens regard themselves and their compatriots as gullible dupes. For example, many Americans seem to have been unwitting participants in Saint Petersburg led strife operations. On November 12, 2016 (four days *after* the presidential election), the IRA used fake social media accounts to induce thousands of people to turn up to simultaneous pro- and anti-Trump rallies in Manhattan. Similar techniques were used to summon Americans to clashing rallies outside a Houston mosque in May 2016.[23] The perceived credibility of democratic citizens is endangered by their apparent readiness to be baited into fighting one another by trolls and automated 'bots' operated by a rivalrous anti-democratic state.

Again, it is important to see that this sort of disinformation is distinctively harmful to democratic societies. Authoritarian regimes may transmit disinformation through state media, but this does not implicate citizens themselves – nor does state legitimacy depend on citizens' individual epistemic reliability. By contrast, democratic citizens who spread or act on social media disinformation become demonstrably complicit. For this reason, the *exposure* of effective social media interference is dangerous, perhaps more dangerous than the disinformation itself. It is a bad thing that

[21] Jeremy B White (2017). 'Facebook says 126 million Americans may have been exposed to Russia-linked US election posts'. *The Independent* Oct 31 2017. https://www.independent.co.uk/news/world/americas/us-politics/facebook-russia-adverts-americans-exposed-trump-us-election-2016-millions-a8028526.html ; Testimony of Sean J. Edgett before the US Senate Judiciary Committee, Subcomittee on Crime and Terrorism. October 31, 2017. https://www.judiciary.senate.gov/imo/media/doc/10-31-17%20Edgett%20Testimony.pdf . p 12. ; Clare Llewellyn, et al. (2018). 'For Whom the Bell Trolls: Troll Behaviour in the Twitter Brexit Debate'. *arXiv* https://arxiv.org/abs/1801.08754 p. 7 ; Olivia Solon (2018). 'Facebook removes 652 fake accounts and pages meant to influence world politics'. *The Guardian* Aug 22 2018. https://www.theguardian.com/technology/2018/aug/21/facebook-pages-accounts-removed-russia-iran

[22] Regina Rini (2017). 'Fake News and Partisan Epistemology'. *Kennedy Institute of Ethics Journal* 27(S2): 43-64.

[23] *United States v. IRA*, 23.; Caroline Haskins (2017). 'I Unknowingly Went to a Trump Protest Organized by Russian Agents'. *Motherboard* Nov 22, 2017. https://motherboard.vice.com/en_us/article/ywb9kx/nyc-trump-election-protest-hack-russian-agents-trolls-government ; Claire Allbright (2017), 'A Russian Facebook page organized a protest in Texas. A different Russian page launched the counterprotest'. *Texas Tribune* November 1, 2017. https://www.texastribune.org/2017/11/01/russian-facebook-page-organized-protest-texas-different-russian-page-l/

citizens unwittingly fall for lies. It is far worse for democratic legitimacy that they come to *know* they and their compatriots have been made fools. It is possible that Russian social media interference operations were *intended* to be exposed, precisely to trigger this effect.[24] This makes disinformation of this form tactically different from traditional propaganda and has important implications for how democratic governments confront the challenge.

Policy recommendations

In this final section I briefly present recommendations for how NATO member states (and other democratic governments) should respond to the likelihood of further, and increasingly more sophisticated, social media disinformation operations.

Given that exposure concentrates the harmful effects of this sort of disinformation, the most important lesson is that **prevention is *much* better than cure**. Once an interference operation has already taken place, exposing it may simply release the poison into democratic circulation. Instead, focus must be on preventing these operations from beginning in the first place, by blocking the creation of fake accounts and limiting the retransmission of false information in social media news feeds. This will obviously require the cooperation of social media companies. Democratic governments should support technical research on algorithmic techniques to pre-emptively identify deceptive social media activity.

Similarly, democratic governments should work to **advance citizen media literacy in ways that do not simply expose the effectiveness of foreign interference**. It is not enough simply to make citizens aware that they and their compatriots have been duped by disinformation, since this awareness is exactly what erodes democratic legitimacy. Rather, media literacy should be focused on helping citizens to learn active techniques for gaining power over disinformation. When there is evidence of success, this should be emphasized so that citizens can recognize their own increasing competence. A good place to start may be publicizing statistics on the large number of citizens who were exposed to IRA disinformation but *did not* share or retweet it.

One project (run by the Alliance for Securing Democracy, with many connections to NATO policy-makers) that may need adjustment is the Hamilton 68 web dashboard, which tracks and publicizes data on active Russian social media interference operations. While this is a highly valuable resource, its public-facing presentation should be more directly integrated with proactive media literacy resources, so that citizens don't simply take away the message that there are very many effective vectors of disinformation duping their compatriots.[25]

Finally, it is essential that **democratic governments, especially NATO member states, cooperate in regulation of social media companies** and provide incentives for them to be transparent. Since the dominant social media firms tend to be based in the United States, this means that the American government must take particular care to govern these firms on behalf of its NATO partners. So far this has not always happened. For instance, though Facebook typically cooperates with American investigative authorities, it tends to stonewall similar demands from countries like Canada and the

[24] I explore this possibility in an essay called 'Weaponized Skepticism', forthcoming in a volume on political epistemology edited by Elizabeth Edenberg and Michael Hannon.

[25] The Hamilton 68 Dashboard is available at https://dashboard.securingdemocracy.org/ . For an example of reporting on Hamilton 68 that emphasizes vulnerabilities, see Tim Mak (2018), 'Tracking Shows Russian Meddling Efforts Ahead of 2018 Midterms', *NPR* February 8, 2018. https://www.npr.org/2018/02/08/584144083/tracking-shows-russian-meddling-efforts-evolving-ahead-of-2018-midterms

UK.[26] The American government must use its authority to compel social media firms to cooperate with legitimate requests from democratic allies, in order to ensure transparency and efficacy in combatting disinformation. These requests must be understood not simply as law enforcement or corporate policy matters, but as defense against existential security threats to all democratic NATO allies.

[26] See Catharine Tunney (2019), 'Facebook's Zuckerberg, Sandberg won't appear before committee, could be found in contempt of Parliament' *CBC News* May 27, 2019. https://www.cbc.ca/news/politics/facebook-contempt-parliament-1.5145347 ; Jane Wakefield (2019), 'Facebook needs regulation as Zuckerbeg 'fails' – UK MPs' *BBC News* February 18, 2019. https://www.bbc.com/news/technology-47255380

Disinformation and Democracy in Historical Perspective

Dr. Heidi Tworek, *Assistant Professor of International History, University of British Columbia*

The last few years have been full of debates about disinformation. Much of the debate was initially about labelling. Scholars soon pushed back against the term "fake news" because it implied that there was such a thing as "true news." The term also offered a weapon to criticize the news industry for correcting genuine mistakes and for leaders simply to decry news that they did not like. After jettisoning the term "fake news," scholars were left with a linguistic challenge of describing this seemingly new phenomenon. Should we differentiate falsified information based upon the intent of the creator? What if someone unwittingly disseminated tweets created by a Russian troll posing as a "real American"? What if the news was accurate but seemed problematic because of how it was spread or who wrote it?

Claire Wardle and Hossein Derakhshan of First Draft News offered a definition of disinformation as "information that is false, and the person who is disseminating it knows it is false. It is a deliberate, intentional lie, and points to people being actively disinformed by malicious actors."[27] Wardle and Derakhshan have created a more detailed taxonomy of what they call "information disorder." Alongside disinformation, they also suggest the terms "misinformation" and "mal-information." Misinformation is false information spread by someone who does not know it is false. Malinformation is true but used to harm a person, country, or party. For example, some leaks or revelations about a person's personal life could harm them, but were accurate. These definitional debates are not arcane, because there are different methods to address different types of problematic content.

More broadly, debates about disinformation cut to the heart of how we understand democratic discourse in a digital age. For years, the Internet's openness appeared to be its strength. Now, many implicitly decry that openness as a weakness that can facilitate foreign interference. A different approach is to consider the historical relationship between disinformation and democracy. Just as democracies value freedom of expression, they have also contended constantly with disinformation (and even sometimes spread it themselves).

Throughout the last few years, multiple media historians have tried to point out the value of historical thinking in our current crisis. Michael Schudson and Barbie Zelizer noted in 2018 that "to act as if today's fake news environment is fundamentally different from that of earlier times misreads how entrenched fake news and broader attitudes toward fakery have been."[28] Rather than focus on this moment's purportedly unprecedented nature, debates about disinformation offer a chance to rethink the history of our democratic discourse and its future. Here, I draw from history to suggest three new approaches to disinformation and digital democracy. First, soft power and hard

[27] Claire Wardle and Hossein Derakhshan, "Thinking about 'information disorder': formats of misinformation, disinformation, and mal-information," in *Journalism, 'Fake News' and Disinformation* (UNESCO, 2018), 43. https://bit.ly/2MuELY5

[28] Michael Schudson and Barbie Zelizer, "Fake News in Context," in *Understanding and Addressing the Disinformation Ecosystem* https://firstdraftnews.org/wp-content/uploads/2018/03/The-Disinformation-Ecosystem-20180207-v2.pdf p. 2.

power have always been intertwined. Second, each new communications technology offers different opportunities to spread disinformation. Finally, disinformation is not a problem we can solve forever. Instead, we can take it as a challenge to reform our democratic discourse for the future.

First, soft power and hard power have always been intertwined. In the 1990s, Joseph Nye differentiated between different strands of power in the international realm. He suggested that "hard power" referred to a country's military and economic might. Meanwhile, "soft power" meant a country's cultural influence, such as the global attraction of Hollywood.[29] Media were generally seen as part of soft power. The concept of soft power became so ubiquitous that Portland, a PR firm, and the USC Center for Public Diplomacy created a soft power index in 2015.[30] Called Soft Power 30, the index uses surveys to measure perceptions of a country abroad alongside quantitative measures such as Internet access.

But soft and hard power were never as separate as Nye's neat division implied. This was particularly true for news. This seemed revelatory with the US election in 2016 or the Brexit referendum. Other scholars have suggested the notion of "sharp power," meaning when authoritarian states take advantage of democracies' openness to spread disinformation, deepen divisions, and generate confusion.[31]

Whether we call it "sharp power" or something else, states have long seen news as part of international relations. My recent book explores one example of this phenomenon, showing why German elites came to believe that news was a cornerstone of political, economic, military, and cultural power at home and abroad in the first half of the twentieth century.[32] I focus less on newspapers than the networks behind the news, specifically one type of media business: the news agency. News agencies emerged in the mid-nineteenth century almost simultaneously to submarine cables. These firms provided most newspapers with their international and national news. As it proved so costly to collect and send international news, only a few news agencies were created. This included the British Reuters (still a major company today) and the German Wolff's Telegraphisches Bureau that initially focused on Europe. To reduce costs, the agencies created a cartel to exchange news that each had gathered.

Around 1900, Germans became increasingly dissatisfied with this system. German political and economic goals had changed. So did their ambitions for news. Many increasingly saw their country as a colonial and aspiring global power. They sought news agencies to bolster that status and increasingly thought that the British and French were using news agencies to malign Germany in places like Latin America that news from Germany never reached directly. Many industrialists thought that they would export more products if more German news were sent to a particular region. Starting around 1900, a news agency consensus emerged among German elites—a belief that news agencies were not simply media businesses, but could achieve broader political, economic, and cultural goals. German elites often disagreed about how to control news agencies or what political and economic goals their news should achieve. But they agreed that news played a central role in public life and international relations.

German news agencies were surprisingly successful in unexpected places. From 1915 to 1917, for example, one news agency (Transocean) sent news from Germany to the United States. The news was written in English and provided German perspectives on the war,

[29] Joseph S. Nye, *Soft Power: The Means to Success in World Politics* (New York: Public Affairs, 2004).
[30] http://www.softpower30.com/
[31] Christopher Walker, "What is 'Sharp Power?'," *Journal of Democracy* 29, no. 3 (2018): 9-23.
[32] Heidi Tworek, *News from Germany: The Competition to Control World Communications, 1900-1945* (Cambridge, MA: Harvard University Press, 2019).

particularly on the Eastern Front. The German government subsidized the agency and aimed to keep the United States neutral during the war. Transocean's articles were widely printed. This did not stop the United States from entering the war in April 1917. But it illustrates that attempts to influence foreign publics through news is nothing new. Information and disinformation have long played a role in statecraft. The more interesting question is why states invest in information to achieve global goals at particular moments, just as Germany did from 1900. Questions about information can be questions about international relations.

A second point from historical analysis is that the Internet is the latest in a long line of communications technologies that offer new methods to spread disinformation. Media scholars talk about "affordances," meaning what a communications technology allows us to do. Each new communications technology offers different affordances, or opportunities, to spread information and disinformation. Radio, for example, was a key boundary-crossing communications technology just under a century ago. Radio jamming and "black radio" were integral features of World War II. Many public media corporations that we praise today, like the BBC, were massively bolstered by investments to fight Nazi and fascist Italian radio sent over international airwaves in the 1930s.

Similarly today, many nation-states use the affordances of social media to try to spread their ideology or to undermine democracies. Russia seemed at the forefront of these developments in 2016. Other nation-states have engaged in similar behaviour since. In the first eight months of 2019 alone, Facebook took down accounts originating in Saudi Arabia, Iran, the Philippines, UAE, Egypt, Thailand, Ukraine, Honduras, and many other places for "coordinated, inauthentic behavior."[33] The difference today is not that nation-states use information networks to try to influence people residing in other states, perhaps particularly democracies. Rather, the difference is the low barrier to entry. Much smaller groups can easily exploit digital networks at far lower cost than a century ago. This creates new aspects of this challenge for social media networks and governments, but the challenge is different in magnitude, not order.

Finally, history reminds us that disinformation is not a problem we can eliminate. Instead, we can take it as a challenge to reform our democratic discourse for the future to include more voices and to encourage new forms of free, fair, and full debate online and offline. As a historian of media and international relations, I can warn that disinformation is important but we must be wary of making it the only issue. Democracy relies on vibrant discourse. Some of that discourse may be disinformation. It can potentially have dramatic effects, such as spreading adherence to anti-vaxxer beliefs, which in turn can cause deaths in the real world from measles. Yet, it is also important to differentiate the potential effects of disinformation. It can be widespread but comparatively ineffective in changing votes. Disinformation may cement someone's beliefs rather than fundamentally alter them. Disinformation may also emerge from domestic as well as foreign actors. Much German disinformation was directed at Germans themselves. More broadly, disinformation becomes more enticing in a world with broader economic, social, and political discontent that people experience in their everyday lives.

Historians love to remind others that every phenomenon has a past. That can help to temper our fears. It can push us to be more accurate about what exactly is new and what is a risk inherent in democracy. The history of disinformation further tempers a subtle nostalgia that has crept into current debates, an idea that we once lived in a utopian world of

[33] https://www.reuters.com/article/us-facebook-saudi-takedowns-factbox/facebook-takedowns-of-coordinated-inauthentic-behavior-in-2019-idUSKCN1UR529

"true news" and fulsome, accurate information. Understanding that disinformation has a history should not create complacency. Rather, it shows that we are paying a heavy price for forgetting about disinformation in the few decades between 1989 and 2016.

When Disinformation Becomes Deadly: The Case of Missing and Murdered Indigenous Women and Girls in Canadian Media

Elisha Corbett, *PhD candidate, Queen's University, and Senior Researcher, National Inquiry into Missing and Murdered Indigenous Women and Girls*

On June 3, 2019, The National Inquiry into Missing and Murdered Indigenous Women and Girls concluded that the disproportionate and distressing deaths and disappearances of Indigenous women and girls is a Canadian genocide.[34] The systemic institutional and cultural causes of this genocide are well documented in the Inquiry's final report; media representation is one of the factors that creates and maintains the violence Indigenous women and girls experience.[35] Indigenous women and girls are misrepresented and underrepresented in the media. Media representation is important because it can influence people's attitudes and opinions, especially in cases where individuals have little first-hand knowledge of the group or issue being covered, as is the case for most Canadians vis-à-vis Indigenous people and issues.[36] The current media misrepresentation of missing and murdered Indigenous women and girls create a narrative that Indigenous women are less deserving of sympathy than other victims of violence. When non-Indigenous Canadians have little sympathy for the plight of Indigenous women and girls, they do not put pressure on the government to help solve this epidemic.[37] Disinformation about missing and murdered Indigenous women and girls has serious consequences: it legitimizes the violence they experience. However, with the advancement of new media, particularly social media, Indigenous people and allies are reframing the narrative around missing and murdered Indigenous women and girls to**,** prevent misinformation and disinformation.

While disinformation is a relatively new concept in political spheres, disinformation about Indigenous women and girls is a part of North-America's colonial history. European settlers have created false images of Indigenous women since contact. When settlers first encountered North-America, Indigenous women were presented as "The Queen." The Queen was militant and mothering.[38] She was depicted in pictures as being draped in leaves, jewelry, and animal skins; she symbolized the beauty of the "new world."[39] However, once settlers desired to conquest more land, the powerful Queen figure was replaced with the "Indian Princess." The Indian Princess was a

[34] Reclaiming Power and Place: The Final Report of the National Inquiry into Missing and Murdered Indigenous Women and Girls. *Supplementary Report: A Legal Analysis of Genocide*, 1. https://www.mmiwg-ffada.ca/wp-content/uploads/2019/06/Supplementary-Report_Genocide.pdf

[35] Reclaiming Power and Place: The Final Report of the National Inquiry into Missing and Murdered Indigenous Women and Girls. *Final Report Volume 1a*, "Deeper Dive: Media Representation", 385. https://www.mmiwg-ffada.ca/wp-content/uploads/2019/06/Final_Report_Vol_1a-1.pdf

[36] Environics Institute. Canadian Public Opinion About Aboriginal Issues in Canada 2016, 13-15. https://www.environicsinstitute.org/docs/default-source/project-documents/public-opinion-about-aboriginal-issues-in-canada-2016/final-report.pdf?sfvrsn=30587aca_2

[37] Amnesty International. *Stolen Sisters*. 2004. 18. https://www.amnesty.ca/sites/default/files/amr200032004enstolensisters.pdf

[38] Green, Rayna. "The Pocahontas Perplex: The Image of Indian Women in American Culture", *The Massachusetts Review* 16, no. 4 (1975).

[39] Ibid.

girlish sexual figure who was often portrayed as scantily clad.[40] She was depicted as cooperating with settlers to colonize land; she symbolized virgin land that was open for consumption.[41] When Indigenous people began to fully resist colonization, the Indian Princess trope was replaced with the squaw. The word squaw means dirty, immoral, and unworthy.[42] As squaw, Indigenous women were the antithesis to the Victorian archetype of womanhood. If Indigenous women were represented as squaw, and Indigenous people more broadly as "savages," colonization could be morally justified in the eyes of the Canadian government.[43]

A clear pattern emerges from examining early media representations of Indigenous women and girls: they are overtly sexual and tied to the process of colonialization. The early representations of Indigenous women are, in part, a deliberate attempt to colonize and conquest Indigenous people and land and are not historically accurate; they were purposely falsified. The early disinformation about Indigenous women and girls has severe consequences as it has and continues to legitimize many forms of violence against them. For example, the squaw stereotype, which presents Indigenous women as inherently sexually available, excuses the sexual violence of white-settler men towards them[44]; it is the age-old false cliché "she was asking for it." Similarly, if Indigenous women are seen as unfit to raise their children in the confines of the Victorian family model because they are viewed as squaw[45], the Canadian government can forcibly remove their children, as was the case during the residential school era, the sixties scoop, and today during the millennial scoop: Indigenous children are still being unlawfully removed from their mothers because they are seen as unfit.

Today, misinformation and disinformation about Indigenous women and girls occurs in news media through its framing. Media frames provide an overarching perspective or narrative to a story through selective emphasis of certain facts or angles.[46] News items are not simply selected, as gatekeeping theory suggests[47]; instead, they are constructed.[48] Misinformation and disinformation are produced both intentionally and unintentionally by journalists. Time-pressures and space limitations that journalists face determine which stories get told and how much detail they will receive.[49] These time constraints often result in simplified and partially told stories, leading journalists to employ stereotypes and tropes for added "colour" and context.[50] However, news stories are nonetheless based on values that society has "created and condoned."[51] Thus, when Canadians accept stories based on disinformation about missing and murdered Indigenous women and girls, it points to a more significant issue about how Canadians view Indigenous women more broadly.

---

[40] Ibid.

[41] Green, Rayna. "The Pocahontas Perplex: The Image of Indian Women in American Culture", *The Massachusetts Review* 16, no. 4 (1975).

[42] LaRocque, Emma. "Métis and Feminist." In *Making Space for Indigenous Feminism*, ed. Joyce Green. Fernwood Publishing, 2018.

[43] Acoose, Janice (Misko-Kisikawihkwe). *Iskwewak. Kah'Ki Yaw Ni Wahkomakanak: Neither Indian Princesses nor Easy Squaws*. Toronto, ON: Women's Press, 1995; Carter, Sarah. *Capturing Women: The Manipulation of Cultural Imagery in Canada's Prairie West*. Montreal, QC: McGill-Queen's University Press, 1997.

[44] Ibid.

[45] Acoose, Janice (Misko-Kisikawihkwe). *Iskwewak. Kah'Ki Yaw Ni Wahkomakanak: Neither Indian Princesses nor Easy Squaws*. Toronto, ON: Women's Press, 1995.

[46] Entman, Robert. "Framing: Toward a Clarification of a Fractured Paradigm." *Journal of Communication*, 43 no. 4 (1993).

[47] Gatekeeping theory states that the media selects and filters information into news stories.

[48] Tolley, Erin. Framed: *Media and the Coverage of Race in Canadian Politics*. UBC Press, 2016.

[49] Ibid.

[50] Ibid.

[51] Ibid.

Non-Indigenous missing and murdered women are framed more compassionately than Indigenous missing and murdered women. The media highlights non-Indigenous women's personalities, families, ambitions, and hobbies, whereas the details of Indigenous women's lives are scant in comparison.[52] Given that articles about Indigenous women are significantly shorter on average than white women, the media does not convey who these women are and what they mean to their families and communities in the same way it does for white victims of violence. [53]

Most often, news media emphasize Indigenous women's and girls' criminal behavior. Indigenous women are framed primarily as sex-workers and criminals. From this framing, a narrative emerges that that Indigenous women are to blame for the violence against them because they engage in "high-risk" lifestyles.[54] The media continually refers to sex-work as a "lifestyle," suggesting that sex-work is an individual choice, even though many Indigenous women might not have other employment opportunities available to them.[55] Media discourse around sex-work as an individual choice suggests that Indigenous women who engage in sex-work and experience violence, as a result, are at fault: by choosing to engage in a "high-risk" lifestyle, Indigenous sex-workers must also accept the consequences of that lifestyle.[56]

The repetitive representation of Indigenous women engaging in "high-risk" lifestyles normalizes the violence against them. In emphasizing Indigenous women's criminal activity in news media, there is no attention to Canada's colonial history that constrains and shapes some Indigenous women's and girls' experiences and opportunities.[57] The violence against Indigenous women and girls is justified because the media's framing signals to the Canadian public that violence against them is not important. The silencing of violence against Indigenous women and girls is made worse in comparison to the media's compassionate framing of white women.

Even Indigenous women who do not engage in "high-risk" lifestyles are framed as "high-risk" individuals by the news media. In 2004, Daleen Kay Bosse, a mother, wife, and student from Onion Lake Cree Nation, went missing. While some news sources mentioned who Daleen was and what she meant to her community, others heavily focused on the time that she went missing, 2:20 am.[58] Continually referring to the time that Daleen went missing

---

[52] Gilchrist, Kristen. "'Newsworthy' Victims? Exploring Differences in Canadian Local Press Coverage of Missing/Murdered Aboriginal and White Women. *"Feminist Media Studies*.

[53] Ibid.

[54] Hallgrimsdotttir Kristin Helga, Rachel Philips, and Cecilia Benoit. "Fallen Women and Rescued Girls: Social Stigma and Media Narratives of the Sex Industry in Victoria, B.C. from 1980-2005." *Canadian Review of Sociology*. (2009); Strega, Susan, Caitlin Janzen, Jeannie Morgan, Leslie Brown, Robina Thomas, and Jeannie Carrirére. "Never Innocent Victims: Street Sex Workers in Canadian Print Media", *Violence Against Women*, 20 no. 1 (2014).

[55] Hallgrimsdotttir Kristin Helga, Rachel Philips, and Cecilia Benoit. "Fallen Women and Rescued Girls: Social Stigma and Media Narratives of the Sex Industry in Victoria, B.C. from 1980-2005." *Canadian Review of Sociology*. (2009); García-Del Moral, Paulina. "Representation as a Technology of Violence: On the Representations of the Murders and Disappearances of Aboriginal Women in Canada and Women in Ciudad Juarez", *Canadian Journal of Latin America and Caribbean Studies*, 36 no. 2 (2011); Strega, Susan, Caitlin Janzen, Jeannie Morgan, Leslie Brown, Robina Thomas, and Jeannie Carrirére. "Never Innocent Victims: Street Sex Workers in Canadian Print Media", *Violence Against Women*, 20 no. 1 (2014).

[56] Ibid.

[57] Ibid.

[58] McKenzie, Holly A. "'She was not into Drugs and Partying. She was a Wife and Mother': Media Representations and (re)presentations of Daleen Kay Bosse (Muskego)." In *Torn from Our Midst: Voices of Grief, Healing and Action from the Missing Indigenous Women Conference, 2008*, ed. Brenda Anderson, Wendee Kubik, and Mary Rucklos Hampton. Canadian Plans Research Centre, 2010.

suggests to readers that Daleen was out late partying, even though she was at a community event on her reserve.[59] Further, many news sources reported on false sightings of Daleen, suggesting to the reader that she was not taken, but rather that she wanted to go missing.[60]

While disinformation runs rampant in mainstream and traditional forms of media, new media, particularly social media, can combat disinformation about missing and murdered Indigenous women and girls. White-settler men dominate mainstream news media. Indeed, both women and Indigenous people are disproportionately underrepresented as journalists and in news media's board of directors.[61] More than seventy-five percent of English language national columnists are men, and women hold only one-third of editorial positions.[62] The homogeneity of the media is worse for racial minorities. Minorities, including Indigenous people, represent only 3.4 percent of news staff.[63] In contrast, everyone and anyone in Canada has access to social media, and there are relatively few boundaries on what a person can post. Social media users do not face the same internal and external pressures as journalists.

This is not to say that new media does not spread disinformation about missing and murdered Indigenous women and girls, but it transcends national and international borders, bringing together Indigenous voices. For example, the #MMIWG, created by Sheila North Wilson, the former Grand Chief of Manitoba Keewatinowi Okimakanak, has been used to raise awareness about missing and murdered Indigenous women and girls, create community events, and reshape the mainstream media's narrative across North-America.[64] #MMIWG is used to post about the actual cases of missing and murdered Indigenous women and girls themselves, disrupting mainstream media's silence on the epidemic.[65] #MMIWG also organizes community events to raise awareness about missing and murdered Indigenous women and girls. For example, #MMIWG is used to lead protests and solidarity marches.[66]

Most notably, #MMIWG is used to reframe the narrative about missing and murdered Indigenous women and girls. A study done in 2018 analyzed tweets using the hashtag from September 2016 to September 2017 and found that the majority of tweets were used to tell the stories of the women themselves, who they were and what they meant to their communities, in the same way that the mainstream media does for white women.[67] #MMIWG is disrupting disinformation in the mainstream media to provide truthful representations of missing and murdered Indigenous women and girls.

Social media campaigns are also being used to disrupt disinformation. For example, the online campaign "She Is Indigenous" aims to challenge the negative stereotypes that non-Indigenous Canadians have about Indigenous women and girls from the media to work towards reducing the violence they experience.[68] The campaign showcases "Indigenous women from across Canada telling *their* stories, in *their* communities, with *their* voices" through weekly Facebook and Instagram posts of Indigenous women across the country.[69] The women featured in the campaign are not missing or murdered, disrupting the mainstream media's exclusive focus on missing and murdered Indigenous women and girls. "She Is Indigenous" demonstrates that Indigenous women and girls

[59] Ibid.
[60] Ibid.
[61] Tolley, Erin. Framed: *Media and the Coverage of Race in Canadian Politics*. UBC Press, 2016.
[62] Ibid.
[63] Ibid.
[64] Moeke-Pickering, Taima, Shelia Cote-Meek, and Ann Pegoraro. "Understanding the ways Missing and Murdered Indigenous Women are Framed and Handled by Social Media Users." *Media International Australia*, 169 no.1 (2018).
[65] Ibid.
[66] Ibid.
[67] Ibid.
[68] She Is Indigenous. https://sheisindigenous.ca/
[69] Ibid.

are newsworthy all the time, not just when they are missing or murdered.

Disinformation about Indigenous women and girls is a part of Canadian's colonial past and present, and this disinformation has severe implications. Mainstream media's inaccurate framing of missing and murdered Indigenous women contributes to non-Indigenous Canadian's negative attitudes and opinions about them, legitimizing the violence they experience. Reframing the narrative, disrupting mainstream media, and letting Indigenous people tell their own stories is imperative to combat disinformation about missing and murdered Indigenous women and girls and to work towards ending violence. The path forward for truthful representations of missing and murdered Indigenous women and girls in the mainstream is unchartered territory; enacting the National Inquiry's calls to justice about the media could be one way forward.

Stalking 'Lolcows' and 'Ratkings': DIY gender policing, far-right digilantes, and anti-transgender violence

Abigail Curlew, *Trudeau Scholar and PhD Candidate, Department of Sociology and Anthropology, Carleton University*
Dr. Jeffrey Monaghan, *Assistant Professor, Institute of Criminology and Criminal Justice, Carleton University*

Sitting on the digital periphery is a lesser-known website called Kiwi Farms (KF), a community of far-right trolls with a penchant for stalking, monitoring, and harassing transgender women who engage in wider public conversations online. Though the KF community has managed to evade mainstream scrutiny, countless trans woman know about their existence by virtue of being targeted by their digital vigilantism and pixelated wrath. As a community of far-right digital vigilantes, what cyberhate researcher Emma Jane calls "digilantism",[70] the cultural make-up of KF is explicitly influenced by the stylized troll culture of fringe digital communities such as 4chan and 8chan.

Typical of the character that animates KF's digital antics, their website sign-up page once read, "Autistics will be laughed at. Trannies will be misgendered. People will try to find where you live".[71] One of the central activities taken up by individuals who participate in the KF forums is to monitor what they refer to as "lolcows" and "ratkings", derogatory terms for progressive trans women whose politics they consider absurd and laughable. Extending trolling culture into a domain of digilantism, KF users spend an extraordinary amount of time and effort monitoring trans women, collecting intelligence, highlighting embarrassing information, crowdsourcing disinformation and posting do-it-yourself (DIY) dossiers to their forum boards. Afterwards, KF users engage in a diversity of digilante tactics, such as "name and shame" attacks, achieved largely through vitriolic comments about a person's trans identity, deadnames,[72] pre-transition photos, and engagement in threatening behavior meant to silence and intimidate their mark. As access to online publics in the era of digital democracy become increasingly crucial to contemporary political engagement and democratic will formation, shadowy digilante campaigns emerging from the KF platform continue to produce coordinated efforts of using cyberhate to chill the capacity of participation in the democratic process by evicting transwomen from the social sphere.

So why do shadowy KF groups invest so much collective enmity towards trans women? In large part, KF users see trans identity as a moral failing at the peak of what far-right proponents call "social justice warrior" (SJW) culture. As an aspect of the broader far-right online movement, KF forums are hives for coordinating digilantism against SJWs (often trans persons demanding equal rights and dignity) with the objective of enacting extrajudicial justice on a marginalized group

[70] Emma Jane, "Online misogyny and feminist digilantism," Journal of Media & Cultural Studies 30, no. 3 (2016): 55-72.
[71] Kiwi Farms. "Sign Up." Last Modified September 2, 2019. https://archive.fo/pytWM.

[72] A deadname is the name given to a person at birth and later changed to reflect a person's actual gender identity. It can be used by hostile communities to embarrass, harass, and delegitimize a trans person in public.

they have construed as a social enemy. Tanner and Campana observe that far-right digilantes are often involved in a form of gatekeeping, "which involves enforcing rules whereby only individuals with certain attributes (racial, ethnic, religious, or societal) can be considered full members of a society".[73] In this way, digital vigilantes are informed by a moral community that is highly misogynistic, racist, and ableist – and forums like KF aim to protect these values at the expense of marginalized groups that they see as not passing the test of social membership.[74]

KF's digilantism is specifically meant to police the boundaries of gender identity and expression. Such practices are organized by logics of DIY gender policing which rely on the visibility afforded to us through social media participation in order to exert a form of violence over those who are deemed as transgressive of the far-right's notions of morality. In the context of internet-based publics, digital violence consists of practices that serve to diminish the life experiences of other online users through coercive means. Such violence has material consequences and often serves as a silencing strategy to discourage participation in wider public conversations. Strategies of DIY gender policing are organized under the punitive logics of policing and intelligence work, and thus make use of digital forms of violence to enact their goals. Indeed, Trottier refers to such practices as the weaponization of visibility or the use of disparate forms of data exhaust[75] left lingering across the internet that when brought together can approximate a person's life experiences, associations, beliefs, and sometimes, dirty secrets.[76] In this way, digital violence is often deployed as a method of maintaining current structural and systemic inequalities that determine who gets to speak in public, who gets to demand rights and call for justice, and who has a seat at the democratic table.

With the growing ubiquity of social media platforms (SMP), North America has undergone massive shifts in how social actors communicate and engage in the polity. Couldry and Hepp describe this new communicative regime as "deep mediatization" where the digital sphere has become a major frame from which social actors engage with the development of self and interaction with society.[77] From this perspective, it is essential that institutions understand that SMP constitute a public in-and-of itself. A useful concept for framing SMP is boyd's "mediated publics", defined as, "environments where people can gather publicly through a mediating technology".[78] In the era of social media, mediated publics are an essential space of democratic engagement where users of multiple backgrounds use social media platforms like Facebook and Twitter to engage in political debate about relevant social issues. Social media users typically participate in a wide variety of SMP, and thus often leave a trail of content data (pictures, comments, and personal information) that are easily searchable by hostile social actors and can be copied out of context. These digital trails have given rise to what has become known as doxxing—the revealing of a user's identity through posting

[73] Samuel Tanner and Aurélie Campana, "'Watchful Citizens' and Digital Vigilantism: A Case Study of the Far Right in Quebec," *Global Crime* (2019): 1-21.
[74] Lars Burr and Steffen Jensen, "Introduction: vigilantism and the policing of everyday life in South Africa," *African Studies*, 63, no. 2 (2004): 139-152.
[75] Data exhaust is a concept that emerges from technologist's occupational discourse that points to the constant generation of data from user digital practices. Such disparate streams of data carry the potential to be transformed from "waste material" to profitable, monetized data. From Shoshana Zuboff, "Big Other: Surveillance capitalism and the prospects of an information civilization," *Journal of Information Technology*, 30 (2015): 75-89.
[76] Daniel Trottier, "Digital Vigilantism as Weaponisation of Visibility," *Philos. Technol*, 30, no. 55 (2017): 55-72.
[77] Nick Couldry and Andreas Hepp, The mediated construction of reality (Cambridge: Polity Press, 2016).
[78] danah boyd, "Social Network Sites: Public, Private, or What?," *Knowledge Tree*, 13 (2007), https://www.danah.org/papers/KnowledgeTree.pdf.

DIY dossiers of personal information to a hostile public. Whether aimed at deliberately undoing a person's anonymity or exposing elements of someone's distant past, doxxing is an attack on an individual's identity and integrity. Not only an attempt to embarrass or discredit, doxxing is also a form of digital violence that calls on other users to harass and attack targeted individuals. These digilante campaigns have added to a climate of violence strategically fostered against trans women.

In the current political climate, trans women have been at the receiving end of constant public debate around the legitimacy of extending basic human rights to protect their wellbeing and dignity.[79] This has led to a "climate of hate" that has defined the role of trans women in the wider polity and has become an "enabling environment" that further legitimizes hate against trans women.[80] As of writing this article, Statistics Canada's methods of tracking violence against trans women are still insufficient which may be a result of the category of "gender identity and expression" only recently being added to federal hate crime legislation. However, according to a dated article from Perry and Dyck, "trans people are perhaps the most targeted hate crime victim in Canada".[81] In the dearth of police data on anti-trans violence, Perry and Dyck rely on data from two major studies that took place almost a decade ago; a report from Egale Canada and another from the Trans PULSE Project. In 2011, Egale reported on data from a survey of 3700 LGBTQ students across Canada and reported that 74% of trans students faced verbal harassment and 37% have faced physical harassment.[82] And in 2013, another survey conducted by the Trans PULSE Project published the results of a survey of 433 trans Ontarians and concluded that 98% of respondents experienced at least one instance of transphobia.[83] The Trans PULSE Project is currently undertaking a larger national census on the experiences of transgender people to fill the extensive data gaps, however, thus far, there has been insufficient research into these issues. Using more recent data from the United Kingdom, where police do track hate crimes targeting trans people, BBC News (2019) reported that police data demonstrated a staggering 81% increase in anti-trans hate crimes between the between 2017 and 2018[84]. Furthermore, this data is likely limited as there is a tendency for trans women to avoid the reporting instances of violence to the police for fear of anti-trans discrimination. Though there is a growing body of literature exploring violence against women in digital spaces,[85]

[79] In the Canadian context, this public conversation around the rights and existence of transgender people peaked with the debates and the eventual adaptation of an *Act to Amend the Canadian Human Rights Code and the Criminal Code*, otherwise known as Bill C-16, which added the category of "gender identity or expression" to the criminal code. This legislation effectively added transgender and nonbinary people as a recognizable group within Canada's hate speech legislation. Right-wing ideologs, such as Jordan Peterson, contributed to fearmongering by putting forward the idea that Canadians could be put in jail for misgendering a transgender person. As Rebecca Thorpe demonstrates, this was entirely inaccurate. See, Thorpe, Rebecca. "Bill C-16 – No, its Not about Criminalizing Pronoun Misuse," *Mark S. Bonham Centre for Sexual Diversity Studies*: http://sds.utoronto.ca/blog/bill-c-16-no-its-not-about-criminalizing-pronoun-misuse/.

[80] Barbara Perry and Ryan Scrivens, "A Climate for Hate? An Exploration of the Right-Wing Extremist Landscape in Canada," *Critical Criminology*, 26 (2018): 169-187.

[81] Barbara Perry and D. Ryan Dyck, "'I Don't Know Where it is Safe?': Trans Women's Experiences of Violence," *Critical Criminology*, 22 (2014): 49-63.

[82] "Every Class in Every School," *Egale*, last modified September 2, 2019, https://egale.ca/every-class/.

[83] "Experiences of Transphobia among Trans Ontarians," Trans PULSE Project, last modified September 2, 2019, http://transpulseproject.ca/wp-content/uploads/2013/03/Transphobia-E-Bulletin-6-vFinal-English.pdf.

[84] "Transgender hate crimes recorded by police go up 81%," *BBC News*, last modified September 2, 2019, https://www.bbc.com/news/uk-48756370.

[85] See: Karla Mantilla (2013). Gendertrolling: Misogyny Adapts to New Media Feminist Studies, 39(2), 563-570; Karla Mantilla (2015).

there has been almost no research into how digital violence impacts trans women.

KF are becoming increasingly notorious for their constant misogynistic attacks on women, specifically trans women, women with disabilities, and fat women. According to the *Rationalwiki*, a treasure-trove of folksy Internet history, the forum board was originally created to exclusively harass an autistic trans webcomic artist named Christine Weston Chandler. It was an offshoot of 4chan and 8chan that focused specifically on troll related doxxing.[86] The community eventually blossomed into the toxic dumpster fire it is today—and began to facilitate the monitoring and harassment of marginalized women. Trolls abide by the cultural logics of the "lulz",[87] which use digital media, Internet memes, and a large dose of cruelty and bigotry to take joy and pride in the harassment of others. After doing a multi-year ethnographic research study of trolling cultures, Phillip's (2015) theorized a set of critical features that represent the trolling culture of "the lulz": (1) Behaviors that fail to celebrate the lulz are excluded, (2) the lulz is secondary to the fallout, (3) Often, the lulz are made at the expense of marginalized people; (4) trolls take nothing seriously; (5) trolls value anonymity; and (6) trolls attack anyone who practices online visibility and demonstrates "real life attachments, interests, and vulnerability".[88] So, in addition to identifying as a moral community that values traditional regimes of cisgender identity and patriarchal methods of organizing society, KF digilantes take a great degree of pleasure in stalking and harassing their marks and witnessing the chaotic fallout that afflicts their lives in the aftermath. For KF users, the use of digital violence to inflict mayhem in the lives of trans women is both a political project and gamified activity that is meant to draw the lines of participation in digital publics.

Though the KF lexicon has an abundance of categories they use to sort their marks, two of the most prolific categories on their forum board are "lolcows" and "ratkings". These categories are an apt example of how politically motivated trolls can combine efforts to achieve the lulz with forms of digital vigilantism in order to engage in gatekeeping and boundary maintenance of gender identity and expression. The term lolcow refers to trans women who are read as SJWs and engage in forms of left-leaning political discourse. Such women are identified as users who might exhibit a strong emotional reaction to being doxxed. KF trolls attempt to herd "lolcows" into a satisfying emotional reaction that they can use as fuel to continue harassing their mark. The more a "lolcow" reacts to such abuse, the more intense that abuse will become. A "ratking" refers to trans women who are engaged in social justice politics with a tightknit group of activists. KF trolls understand a ratking to be inauthentically political (aka, "professional victims") and read them as engaging in political discourse in order to accumulate personal influence, fame, and monetary gain. According to the *Lolcow wiki*, a KF affiliated website dedicated to tracking troll culture, "a rat king is a nest of rats who become bound together with excrement, blood, hair, or knotted by their own tails".[89] The categories themselves are a form of violence that operate to dehumanize and delegitimize the political work of transgender activists, journalists, and scholars and thus compromise our ability to engage in digital democracy.

Gendertrolling: How Misogyny went viral. Santa Barbara: Praeger; Emma A. Jane. (2017). Misogyny Online: A short (and brutish) history. London: Sage Publishing.

[86] "Kiwi Farms," *Rational Wiki*, last modified September 2, 2019, https://rationalwiki.org/wiki/Kiwi_Farms.

[87] An acronym referring to lol, or laugh our loud, that was adopted into trolling vernacular in the early years of 4chan.

[88] Whitney Phillips, This Is Why We Can't Have Nice Things: Mapping the Relationship between Online Trolling and Mainstream Culture (Massachusetts: The MIT Press, 2015).

[89] "Rat King," *Lolcow Wiki*, last modified September 2, 2019, https://lolcow.wiki/wiki/Rat_King#cite_note-wp-rat-king-1.

As Phillips noted in the wake of her ethnography on trolling, social media and the ability to engage in pseudonymous or anonymous communication is not the root cause of digital violence.[90] Troll vigilante practices emerge from an enabling mainstream culture that values sensationalist news media, ongoing political upheaval, and constant debates about the basic human rights and dignity of trans women. Restricting and censoring speech over social media platforms not only ignores the root issues, but also impacts the political organizing of marginalized groups. Phillips aptly observes, "attempt to smoke out the trolls, in other words, and you simultaneously smoke out the activists".[91]

Just the same, there is an urgent need for the public to grapple with these emerging forms of digilantism and digital violence across social media platforms. Digital communities like KF have produced new cultures of punishment often aimed towards spectacular forms of shaming, embarrassment, and social death. Emerging forms of digital punitiveness are resulting in increasingly hostile and reactionary cultures that are having imminent consequences for the right to equal participation in digital democracies. What's more is that digilantism is a form of digital politics that strives to translate into material violence by fomenting a culture of hate paralleled with the doxxing and harassment of trans women. Calls for violence are especially pronounced among far-right actors. While forms of digilante practices may be taken by a spectrum of political actors, far-right digilantes are far more comfortable and politically-acclimatized towards acts of violence.

Though potential solutions to these emerging vectors of targeted violence are still uncertain, our research suggests that an increase in state surveillance and criminalization may not be effective approaches. A punitive response would very likely result in a doubling down effect in far-right digilante communities and would almost certainly increase a climate of fear and anxiety that could likely further dampen democratic participation. Deplatforming has some immediate impacts for stopping digilante attacks yet protecting the rights of transwomen requires far more action than attempts to police the internet. In order to curb the rise in digilantism, the general public and policy makers need to acknowledge that these groups are an outgrowth of racist and misogynistic cultural venues. In addition to steps like deplatforming, far greater leadership is required in cementing trans rights across the North American political landscape, broadening and affirming critical supports for trans women suffering from targeted violence.

[90] Whitney Phillips, This Is Why We Can't Have Nice Things: Mapping the Relationship between Online Trolling and Mainstream Culture (Massachusetts: The MIT Press, 2015).

[91] Ibid, 155.

From a Sleazy Reddit Post to a National Security Threat: A closer look at the deepfake discourse

Victoria Heath, *MGA, Communications Manager for Creative Commons and Senior Research Fellow, NATO Association of Canada*

"Deepfake" is Everywhere

The term is featured in tech and national security blogs; sprawled across nightly news' tickers; discussed on expert panels in Davos; and whispered about behind closed doors at the Pentagon.

Political leaders, terrified at how quickly the technology behind deepfakes is developing (i.e. deep learning), are clamouring to introduce legislation to regulate it. Just under a year ago, the United States (U.S.) Senator Ben Sasse introduced a bill titled, *Malicious Deep Fake Prohibition Act of 2018*—reflecting the nervous energy engulfing both local and national governments.[92] Just before introducing the bill, Sen. Sasse wrote in the Washington Post, "Deepfakes...are likely to send American politics into a tailspin, and Washington isn't paying nearly enough attention to the very real danger that's right around the corner."[93]

According to media reports, however, security researchers and military agencies are paying attention. The U.S.' Defence Advanced Research Projects Agency (DARPA), for example, is spending millions on media forensics to find a "technological solution for spotting manipulated videos."[94] In Canada, the Communications Security Establishment (CSE) warns, "Evolving technology underpinned by AI, such as deep fakes, will almost certainly allow threat actors to become more agile and effective when creating false or misleading content intended to influence voters, and make foreign cyber interference activity more difficult to detect and mitigate."[95]

In more alarming language, Professors Robert Chesney and Danielle K. Citron write:

> The array of potential harms that deep fakes could entail is stunning. A well-timed and thoughtfully scripted deep fake or series of deep fakes could tip an election, spark violence in a city primed for civil unrest, bolster insurgent narratives about an enemy's supposed atrocities, or exacerbate political divisions in a society. The opportunities for the sabotage of rivals are legion—for example, sinking a trade deal by slipping to a foreign leader a deep fake

[92] Malicious Deep Fake Prohibition Act of 2018, S.3805, 115th Congress. (2018). Retrieved from: https://www.congress.gov/bill/115th-congress/senate-bill/3805/text?format=txt

[93] Sasse, B. (2018, October 19). *This new technology could send American politics into a tailspin*. Washington Post. Retrieved from: https://www.washingtonpost.com/opinions/the-real-scary-news-about-deepfakes/2018/10/19/6238c3ce-d176-11e8-83d6-291fcead2ab1_story.html

[94] Kampf, Stephanie; Kelley, Mark. (2018, November 18). *A new 'arms race': How the U.S. military is spending millions to fight fake images*. CBC. Retrieved from: https://www.cbc.ca/news/technology/fighting-fake-images-military-1.4905775

[95] (2019). *Cyber Threats to Canada's Democratic Process*. Communications Security Establishment, Government of Canada. 18. Retrieved from: https://cyber.gc.ca/sites/default/files/publications/tdp-2019-report_e.pdf

> purporting to reveal the insulting true beliefs or intentions of U.S. officials.[96]

Even social media giants are recognizing there may be a problem. In response to pressure from civil society groups and lawmakers, Facebook recently teamed up with Microsoft to launch the $10 million USD "Deepfake Detection Challenge."[97] This Challenge is meant to "produce technology that everyone can use to better detect when AI has been used to alter a video in order to mislead the viewer."[98]

Therefore, from a bird's eye view, the danger from deepfakes to our political system and national security may seem imminent and profoundly concerning—and it is, arguably. However, we must look beyond the current fervor to see the more immediate and pressing danger of deepfakes: bullying, harassment, and violence towards individuals, particularly those who have been historically discriminated against or are vulnerable.

The Rise of the Deepfake Discourse

With all of this attention on deepfakes today, it's easy to forget that this term is new. In fact, its known origins date as late as December 2017, when an anonymous Reddit user who called themselves "deepfakes" began posting "digitally superimposed faces of celebrities on actors in pornographic content" through deep learning algorithms.[99] Following an initial report by Motherboard in 2017, the term exploded across the web and concern for the threat from deepfakes quickly grew.[100] This is evident in an early February 2018 article published by The Outline, which quotes several attendees at the DARPA Media Forensics program meeting feeling "blindsided" by the Reddit deepfakes.[101] Hany Farid, a professor of computer science at Darmouth University, told The Outline, "[the] nightmare situation of somebody creating a video of Trump saying, 'I've launched nuclear weapons against North Korea,' and that video goes viral, and before anyone gets around to realizing it's fake, we have a full-blown nucleur holocaust...I think we can agree that's not entirely out of the question right now."[102]

An article published just a few days later on Lawfare lists several nightmare scenarios that could result from the release of a deepfake video.[103] However, the authors also explain the specific threats to individuals. Beyond the harmful and despicable "conscription of individuals" into "fake porn scenarios," they write, "We can expect to see deep fakes used in other abusive, individually-targeted ways, such as undermining a rival's relationship with fake

---

[96] Chesney, Robert; Citron, Danielle K. (2018, October 16). *Disinformation on Steroids: The Threat of Deep Fakes*. Council on Foreign Relations. Retrieved from: https://www.cfr.org/report/deep-fake-disinformation-steroids

[97] The Deepfake Detection Challenge. Retrieved from: https://deepfakedetectionchallenge.ai/index.html

[98] Schroepfer, Mike. (2019, September 05). *Creating a data set and a challenge for deepfakes*. Facebook. Retrieved from: https://ai.facebook.com/blog/deepfake-detection-challenge/

[99] Van de Weghe, Tom. (2019, May 29). *Six lessons from my deepfakes research at Stanford*. Medium. Retrieved from: https://medium.com/jsk-class-of-2019/six-lessons-from-my-deepfake-research-at-stanford-1666594a8e50

[100] Cole, Samantha. (2017, December 11). *AI-Assisted Fake Porn Is Here and We're All Fucked*. VICE Motherboard. Retrieved from: https://www.vice.com/en_us/article/gydydm/gal-gadot-fake-ai-porn

[101] Christian, Jon. (2018, February 01). *Experts fear face swapping tech could start an international showdown.* The Outline. Retrieved from: https://theoutline.com/post/3179/deepfake-videos-are-freaking-experts-out?zd=2&zi=qzsymyot

[102] Ibid.

[103] Chesney, Robert; Citron, Danielle. (2018, February 21). *Deep Fakes: A Looming Crisis for National Security, Democracy and Privacy?* Lawfare. Retrieved from: https://www.lawfareblog.com/deep-fakes-looming-crisis-national-security-democracy-and-privacy

evidence of an affair or an enemy's career with fake evidence of a racist comment."[104]

Over time the threat to individuals from deepfakes faded from the headlines and broader threats to democracy and national security took over. However, those fears have largely proven to be unfounded—at least for now. Russel Brandom, a journalist at The Verge argues, "...increasingly, the panic around AI-assisted propaganda seems like a false alarm." He continues, "There's still real damage being done by deepfake techniques, but it's happening in pornography, not politics...But most deepfake coverage has treated pornography as an embarrassing sideshow to protecting the political discourse."[105]

So Far, Most Deepfakes Threaten Individuals; Not Nations

Brandom's critique of the discourse on deepfakes was arguably an outlier at the time. However, these issues are finding their way back into the headlines—thanks in part to Deeptrace, a technology company which recently found that of the estimated 14,678 identifiable deepfakes online, 96% were non consensual pornographic content and at least 99% were of women who work in the entertainment industry (e.g. celebrities, etc.).[106] This led CNN to report, "While much of the coverage about deepfakes has focused on its potential to be a tool for information warfare in politics, the Deeptrace findings show the more immediate issue is porn.."[107]

Tom Van de Weghe, John S. Knight Fellow at Stanford University, writes that eventually anyone can be a target of deepfakes due to the democratization of the required technology and the increase of "photo source data material."[108] "Unfortunately," Weghe explains, "the most imminent threat of deepfakes comes from weaponizing them against women. Deepfake creators use their faces on pornographic content without consent. This trend is better known as *revenge porn* and represents a degrading way of humiliating, harassing and abusing victims."[109]

Sam Gregory, Program Director at WITNESS, tells Wired that there are also instances in which deepfakes have been used to "harass or discredit women journalists or activists," as well as female politicians.[110] This was evident in the doctored viral video of U.S. House Speaker Nancy Pelosi released in May 2019 that portrayed her as slurring her speech—eerily similar to the doctored videos that went viral in 2016 of then U.S. Presidential candidate Hillary Clinton, edited to make her look impaired and sickly.[111] It's important to note that these videos were not only a political attack, but a personal attack on these female politician's ages, abilities, and gender. These examples are indicative of what's to come regarding deepfakes and reflect larger trends in the digital space, as Brandom writes, "the deepfake story is about misogynist harassment rather than

---

[104] Ibid.

[105] Brandom, Russell. (2019, March 05). *Deepfake propaganda is not a real problem*. The Verge. Retrieved from: https://www.theverge.com/2019/3/5/18251736/deepfake-propaganda-misinformation-troll-video-hoax

[106] Brown, Jennings. (2019, October 07). *California Bans Deepfakes in Porn and Politics*. Gizmodo. Retrieved from: https://gizmodo.com/california-bans-deepfakes-in-porn-and-politics-1838844251

[107] Metz, Rachel. (2019, October 07). *The number of deepfake videos online is spiking. Most are porn*. CNN. Retrieved from: https://www.cnn.com/2019/10/07/tech/deepfake-videos-increase/index.html

[108] Van de Weghe, Tom. (2019, May 29).

[109] Van de Weghe, Tom. (2019, May 29).

[110] Simonite, Tom. (2019, October 06). *Prepare for the Deepfake Era of Web Video*. Wired. Retrieved from: https://www.wired.com/story/prepare-deepfake-era-web-video/

[111] Kelly, Makena. (2019, May 24). *Distorted Nancy Pelosi videos show platforms aren't ready to fight dirty campaign tricks*. The Verge. Retrieved from: https://www.theverge.com/2019/5/24/18637771/nancy-pelosi-congress-deepfake-video-facebook-twitter-youtube

geopolitical intrigue, with less obvious implications for national politics."[112]

If we zoom out and look at technology-facilitated violence, abuse, and harassment more generally (which includes deepfakes), evidence shows that the "Internet is being used in a broader environment of widespread and systemic structural discrimination and gender based violence against women and girls."[113] According to the UN, 23% of women reported experiencing this form of gender based violence at least once in their life, and as early as 15.[114] In 2018, Amnesty International found that 76% of the 4,000 women they polled "experienced abuse or harassment on a social media platform."[115]

Studies have also found that the most vulnerable women are "human rights defenders, women in politics, journalists, bloggers, young women, women belonging to ethnic minorities and Indigenous women," as well as women belonging to the LGBTQ2+ community and those with disabilities.[116] According to another Amnesty International study, on average, 7.1% of the tweets sent to the women journalists and politicians surveyed were "abusive or problematic." Further, women of color were 34% more likely to be targeted than white women, and black women in particular were targeted at a higher rate.[117]

There is also growing concern regarding the abuse of smart technologies, such as cellphones. For instance, researchers at the Citizen Lab found that domestic violence perpetrators have installed spyware on their target's phones to surveil, harass, and control them—effectively turning the software meant to "facilitate intimate partner surveillance, parent-child monitoring, or monitoring of employees" into "stalkerware."[118] The researchers wrote, "As new technologies have seeped into everyday life, aggressors have adopted and repurposed them to terrorize, control, and manipulate their current and former partners."[119] Recent reports have also focused on the abuse of smart home technologies, such as home security systems. Chantel Nelson, a Toronto-based social worker, told journalist Takara Small earlier this year that she's increasingly counselling domestic abuse victims who have been entrapped in their homes by these systems.[120]

Historically, this form of gender based violence has not been taken seriously. This is due to many factors, including the lack of laws protecting victims and the difficulty in identifying perpetrators, as well as the trivialization of online harassment and "victim blaming." The Global Fund for Women, for instance, reports that in some cases online violence is "laughed off" by the victim's family

[112] Brandom, Russell. (2019, March 05).

[113] Malicious Deep Fake Prohibition Act of 2018, S.3805, 115th Congress. (2018). Retrieved from: https://www.congress.gov/bill/115th-congress/senate-bill/3805/text?format=txt
Report of the Special Rapporteur on violence against women, its causes and consequences on online violence against women and girls from a human rights perspective. A/HRC/38/47, UN General Assembly. (2018, June 18). Retrieved from: https://ap.ohchr.org/documents/dpage_e.aspx?si=A/HRC/38/47

[114] Ibid.

[115] (2018, March 03). *Online abuse of women thrives as Twitter fails to respect women's rights*. Amnesty International. Retrieved from: https://www.amnestyusa.org/reports/online-abuse-of-women-thrives-as-twitter-fails-to-respect-womens-rights/

[116] A/HRC/38/47, UN General Assembly. (2018, June 18).

[117] Amnesty International. (2018, March 03).

[118] Parsons, Christopher; Molnar, Adam; Dalek, Jakub; Knockel, Jeffrey; Kenyon, Miles; Haselton, Bennett; Khoo, Cynthia; Deibert, Ron. (2019, June 12). *The Predator in Your Pocket: A Multidisciplinary Assessment of the Stalkerware Application Industry*. The Citizen Lab, University of Toronto. Retrieved from: https://citizenlab.ca/2019/06/the-predator-in-your-pocket-a-multidisciplinary-assessment-of-the-stalkerware-application-industry/

[119] Ibid.

[120] Small, Takara. (2019, January 09). *How Smart Home Systems & Tech Have Created A New Form Of Abuse*. Refinery 29. Retrieved from: https://www.refinery29.com/en-ca/2019/01/220847/domestic-abuse-violence-harassment-smart-home-monitoring

or friends, and more often for girls, they're Internet access is simply restricted when they report being harassed.[121] "In a world where we seamlessly navigate the online and the offline everyday," explains Bishakha Datta, Executive Director of Point of View, "it is crucial for us to address the violence that women face in both realms."[122] In regards to deepfakes in particular, we must first grapple with the fact that they are the latest, sinister example of online violence.

Informing the Discourse with Feminist Scholarship

It's important to recognize that deepfakes are a threat to both national and personal security—and this reality implores us to understand further what many feminist scholars have been arguing for decades: It's impossible to separate public violence from private violence; state violence from domestic violence. As Dr. Patricia Owens argues, "There is only violence that is *made* 'public' and violence that is *made* 'private'."[123] However, the definition of national security or security more broadly often excludes "'private,' 'domestic,' or 'individual' security issues such as gender-based violence," and unfortunately, the mainstream discourse regarding the threats from deepfakes has also excluded these issues.[124]

To inform the discourse regarding deepfakes, as well as create effective policies and regulations to protect individuals who may be targeted, it's imperative to include feminist security and international relations scholarship—something that the cybersecurity field, in particular, has failed to do. In her recent paper, *Safe at Home: Towards a Feminist Critique of Cybersecurity*, scholar Julia Slupska argues:

> Just as feminist theorists critiqued the public/private binary that shielded violence in the home from outside scrutiny, we now need a feminist critique of cybersecurity. Feminist cybersecurity will ask necessary questions: For whom are these technologies made? Where could interventions take place? What trade-offs are made in technology companies when concerns do arise? Whose security is cybersecurity?[125]

Feminist science and technology studies' (STS) theories should also inform the deepfake discourse. In particular, the social constructivist notion that technology is neither neutral nor static is important to take note of when studying how technology impacts society. As Dr. Judy Wajcman explains, social relations (including gender relations) "are materialised in technology" and "the fate of a technology depends on the social context and cannot simply be read off fixed sets of power arrangements."[126] Put simply, technologies—including deepfakes—both shape society and are shaped by it.

As of now, and in the foreseeable future, the primary targets of deepfakes are women and

[121] (2019). *Online violence: Just because it's virtual doesn't make it any less real*. Global Fund for Women. Retrieved from: https://www.globalfundforwomen.org/online-violence-just-because-its-virtual-doesnt-make-it-any-less-real/

[122] Ibid.

[123] Owens, Patricia. (2008, October 03). Distinctions, distinctions: 'public' and 'private' force?. *International Affairs*, Volume 84, Issue 5, September 2008, Pages 977–990. Retrieved from: https://doi.org/10.1111/j.1468-2346.2008.00750.x

[124] Slupska, Julia. (2019, May 01). Safe at Home: Towards a Feminist Critique of Cybersecurity. *St. Anthony's International Review,* No. 15, 2019, Pages 83-99. Retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3429851

[125] Ibid: 98.

[126] Wajcman, Judy. (2009, January 08). Feminist theories of technology. *Cambridge Journal of Economics*, Volume 34, Issue 1, January 2010, Pages 143–152. Retrieved from: https://doi.org/10.1093/cje/ben057

the motivations aren't necessarily political but rooted in misogyny and discrimination. This is a reflection of the broader trends seen in technology-facilitated violence, abuse, and harassment. Yet the discourse hasn't echoed that reality.

Therefore, as the discourse regarding the threats from deepfakes matures into actual political, legislative, and technical action, it's imperative that we ask the questions and integrate the theories proposed by feminist scholars, such as Owens, Slupska and Wajcman. It's also imperative that we no longer ignore the very real threat deepfakes pose not only to our democracy and national security, but also to ourselves as individuals. If we don't change our discourse, efforts to reduce the threat from deepfakes—or any other emerging technology—may ultimately fail. Even worse, they may reinforce or exacerbate the existing biases, social norms and structures that made said technology harmful in the first place.

Online Disinformation Threats in the 2019 Canadian Federal Election: Who is Behind them and Why?

Christian Piccard, *MA, Researcher,*
Canadian Defense and Security Network

As we have seen in other countries (2016 elections in the USA, the United Kingdom and The Netherlands being the main examples), elections and referendums are prime targets for online disinformation. With Canada entering into its 2019 federal election, now is a good time to reflect on the disinformation threats the Canadian electoral process might be facing. This is particularly important considering that in Canada, online sources are one of the main sources of news for Canadians[127], with around 75 % of people getting their news online, including through social media sources.

There are three main possible sources of disinformation that Canadians should look out for: Russia, China, and Canadian politicians or special interest groups. As we will see further, out of the three, Canadians politicians are the most likely sources of disinformation, while Russia and China will probably be much narrower in their interventions. This can be explained by the issues at stake. On one hand, contrary to the USA or the UK, Canada is a middle power and does not present a major or existential threat to Russia and China. It does not mean there are no issues between them, but they are less directly relevant at the level of the global world order. Hence, there is no real need for China and Russia to engage aggressively and at a large scale in Canada, with operatives able to focus on specific issues or on broad efforts to undermine global faith in democratic institutions. On the other hand, Canadian politics are getting more and more divided, thanks, in part, to the rise of wedge politics[128]. This breed a fertile ground for otherwise marginal political trends to get more clout, especially through social media.

What about Russia?

While many states have a long history of online disinformation warfare (the IDF and its online fights against pro-Palestine online content are famous), Russia has emerged as a pioneer of large-scale online disinformation operations. For Russia, propaganda is part of a broader strategy against Western countries called hybrid warfare[129]. The idea behind this strategy is that since a direct military confrontation against Western countries (mainly NATO members) would be a disaster, indirect confrontation on multiple targets is the preferred method for Moscow to put pressure on its adversaries. This has resulted in various smaller confrontations, such as the 2007 cyberattacks against Estonia, stealth submarines in the Baltic sea, the Crimea annexation, as well as the kidnapping or murder of various targets in Western countries.

One easily sees how propaganda and online disinformation can fit into this picture. Russia

[127] Brin, Colette (Université Laval), "Digital News Report – Canada", for Reuters Institute/University of Oxford, consulted online on Sept. 2019: http://www.digitalnewsreport.org/survey/2018/canada-2018/

[128] Fyfe Toby & Mike Colledge (Aug. 2nd 2019), "COMMENTARY: How the 'wedge solution' is further dividing Canadians", *Global News*, online: https://globalnews.ca/news/5702306/wedge-solutions/

[129] Żaryn, Stanisław (Aug. 9th, 2019), "Russia's hybrid warfare toolkit has more to offer than propaganda", *Defense News*, online: https://www.defensenews.com/opinion/commentary/2019/08/09/russias-hybrid-warfare-toolkit-has-more-to-offer-than-propaganda/

has been maintaining troll farms since at least 2014[130] and has been involved in some high-profile cases in 2016[131] (notably the Netherlands referendum for the Ukraine accession to EU, the Brexit vote, and the US presidential elections). To give an idea of how far Russian propaganda can go, a Congress report revealed that the groups facing off in a 2016 Texas double-rally on immigration[132] were, in fact, mobilized on Facebook by Russian trolls. But what about Canada? As mentioned before, Canada is not a prime target for the control of the international agenda, despite its position as a G-7 member. However, Russia does have two main issues with Canada: Artic sovereignty[133] and Canada's military presence in Latvia[134] and in Ukraine[135].

Arctic sovereignty is an issue that divides many countries, because of the huge potential for access to natural resources and the rapidly emerging trade routes of the Northwest Passage. While we can expect tensions to remain low between Canada and the USA or Denmark, this cannot be said of Russia. For Moscow, the Arctic region is seen as vital to its economic and military interests. In recent years, Russians have been restoring Arctic bases and have sent icebreakers and nuclear submarines into the region to reassess their claims. Since Canada is claiming the Northwest Passage is not an international passage and is advocating a multilateral regime in the Arctic, there is a direct conflict between the interests of Moscow and Ottawa.

On the military side, Canada is involved in direct confrontation with Russia in Latvia and in Ukraine, in each of which Canada is exercising a noticeable military leadership. In Latvia, Canada is commanding a NATO Enhanced Forward Presence (EFP) battle group. The goal of the NATO EFP is to provide assurance and deterrence power to NATO members that feel threatened by Russia. In fact, Latvia is home to Canada biggest military deployment and has even offered to help Canada[136] fight against online disinformation in the federal election. In Ukraine, Canada is part of multilateral training efforts, alongside, among others, the USA and the United Kingdom, to help strengthen Ukrainian military forces in the face of Russian aggressions in Crimea and South Ossetia. Canada has been part of this effort since the beginning, and has even increased its participation over time.
We can easily see why Russia would feel the need to fight Canada on the topics of the Arctic sovereignty and its military presence in Latvia

---

130 Lee, Dave (Feb. 16th, 2018), "The tactics of a Russian troll farm", *BBC News*, online: https://www.bbc.com/news/technology-43093390

131 Stronski, Paul & Richard Sokolsky (Dec. 14th 2017), "The Return of Global Russia: An Analytical Framework", *Carnegie Endowment For International Peace*, online: https://carnegieendowment.org/2017/12/14/return-of-global-russia-analytical-framework-pub-75003

132 Bertrand, Natasha (Nov. 1st, 2017), "Russia organized 2 sides of a Texas protest and encouraged 'bot sides to battle in the streets' ", *Business Insider*, online: https://www.businessinsider.com/russia-trolls-senate-intelligence-committee-hearing-2017-11

133 Blanchfiled, Mike (Sept. 8th, 2019), "Russia Will Meddle in Federal Election to Serve Its Arctic Goals: Study", *Canadian Press/Huffington Post Canada*, online: https://www.huffingtonpost.ca/entry/russia-federal-election-meddling-arctic_ca_5d7562c1e4b0fde50c28934f

134 Kwok, Tiffany (June 4th, 2019), "Preparing Canada to Combat Disinformation in the Upcoming Federal Election", *NATO Association of Canada/Centre for Disinformation Studies*, online: http://natoassociation.ca/preparing-canada-to-combat-disinformation-in-the-upcoming-federal-election/

135 Canada – National Defence (Mar. 18th, 2019), "Canada extends its military training mission in Ukraine", online: https://www.canada.ca/en/department-national-defence/news/2019/03/canada-extends-its-military-training-mission-in-ukraine.html

136 Chase, Steven (May 6ht, 2019), "Latvia offers Canada help in 'likely' election interference", *The Globe and Mail*, online: https://www.theglobeandmail.com/politics/article-latvia-offers-canada-help-in-likely-election-interference/

and Ukraine. In both case, Canada' stance is in direct opposition to Russia's national interests.

What about China?

In the past two years, Canada and China have clashed over three main issues: Canadian agricultural exports to China, the arrest of Huawei's Meng Wanzhou, and Canada's criticisms against China's human rights violations. While the agricultural exports are unlikely to be a matter of a high-profile political importance during the election (they haven't been before), the other issues are.

Recently, Facebook and Twitter announced the suspension of various accounts on their platforms linked to the Chinese authorities. According to the tech giants[137], these accounts were spreading disinformation and propaganda in the Hong Kong protests to the benefit of the Hong Kong government. While Chinese officials dismissed the authorities being behind these accounts, they nonetheless condemn the move by Facebook and Twitter, stating these accounts should be able to express their point of view in the name of free speech. Back to Canada, seeing how China personally targeted Chrystia Freeland[138] for her recent condemnation of its conduct in Hong Kong, it is not far-fetch to believe that some members of the Canadian government may be victims of China-sponsored online smear campaigns during the elections.

But it is the arrest of Huawei's Meng Wanzhou that is the most likely to drive Chinese disinformation campaigns during the election. The arrest of Mrs. Wanzhou by Canadian border officers in Vancouver was the result of a request by American authorities. By acceding to this request, Canada jumped directly into the trade war between China and the USA. Following the arrest of Mrs. Wanzhou, China also retaliated by jailing two Canadians citizens[139], making the whole diplomatic standoff much more acute and leading to various wars of words[140] between China, Canada and the USA. This tense standoff will likely result in China resorting to disinformation against Canada during the election, probably by doubling-down on its existing claims regarding the illegality of Mrs. Wanzhou arrest and the false charges of espionage levelled against the two detained Canadians.

What will be the most likely source of disinformation?

In April 2019, a social media campaign against the Quebec's immigration bill[141] was led by Iranians stuck in an administrative limbo created by this bill. It showed us that foreigners can pick on very specific targets, despite the

---

137 Timberger, Craig, Drew Harwell & Tony Romn (Aug. 20th, 2019), "In accusing China of disinformation, Twitter and Facebook take on a role they've long rejected", *The Washington Post*, online: https://www.washingtonpost.com/technology/2019/08/20/after-twitter-facebook-blame-china-hong-kong-disinformation-government-defends-its-right-online-speech/

138 Vanderklippe, Nathan & Janice Dickson (Aug. 21st, 2019), "China singles out Chrystia Freeland for unusually personal rebuke over her comments about Hong Kong", *The Globe and Mail*, online: https://www.theglobeandmail.com/world/article-a-warning-chinas-central-television-takes-direct-aim-at-canadian/

139 Connolly, Amanda (Aug. 25th, 2019), " 'Name it and shame it' when China acts out, urges State Department spokesperson", *Global News*, online: https://globalnews.ca/news/5806284/state-department-detained-canadians-action/

140 Chase, Steven (Aug. 23rd, 2019), "China accuses Canada and U.S. of 'singing a duet' to confuse world over Meng and Canadian detainees", *The Globe and Mail*, online: https://www.theglobeandmail.com/politics/article-china-accuses-canada-and-us-of-singing-a-duet-to-confuse-world/

141 Schué, Romain (Mar. 1st, 2019), "Un millier d'Iraniens s'activent sur Twitter pour dénoncer la réforme de l'immigration au Québec", *Radio-Canada*, online: https://ici.radio-canada.ca/nouvelle/1156048/immigration-twitter-comptes-mobilisation-internet-legault

language barrier (a mainly French speaking issue targeted by Iranians). That taught us that "under the radar" propaganda geared toward non-English and non-French speaking or ethnic groups in Canada is a real and concerning possibility. In this regard, the country is vulnerable to targeted propaganda by China and Russia because of its diverse ethnic demography. According to StatCan data of the 2016 census[142], there are about 1,3 million people of Ukrainian ancestry, 1 million of Polish ancestry, and 600,000 of Russian ancestry. All these people could be easily targeted by Russian propaganda in various languages with which not all Canadian security and intelligence officers are familiar. The same can be said for China, since about 1,7 million Canadians are of Chinese ancestry. Moreover, about 70 % of Canadians claiming to be of Asian origin are first-generation immigrants, making them more likely to follow Asian media outlets.

Ultimately, by targeting specific issues, foreign powers know they have very few chances of influencing the public through disinformation and propaganda in Canada, but they can at least cast some doubts on the legitimacy of the Canadian elections. That means Canadian politicians (and, first among them, the government) will have to spend more time and resources to reassure the public of their legitimacy. In return, they will have less time and resources to address some other issues affecting Russia and China. That being said, the main issues Canada is facing against Russia and China are not existential ones. It is then highly unlikely these two countries will launch operations of a comparable scale to what we have seen in the USA through the findings of the Mueller report.

Meanwhile, Canada has not been insulated from the rise of wedge politics. If the situation is far from being as acute as what we see in the USA, it has still resulted in some polarized political debates in recent years, with the election of the controversial Conservative leader Doug Ford in Ontario being a prime example. Internal propaganda and disinformation to feed wedge politics will certainly be the main threat facing the Canadian public during the elections, from sources ranging from established politicians to interest groups.

Maxime Bernier's People's Party of Canada is a prime example. The party hosts a number of conspiracy theorists[143] among its candidates and Bernier himself is a declared climate sceptic, in opposition to the widely accepted scientific findings of anthropogenic climate change. While the PPC is not alone in terms of alt- and far-right groups wishing to enter the elections[144] (since this movement is getting more traction in Canada, as VICE's Mack Lamoureux' multiple reports show[145]), mainstream politicians and groups can also resort to propaganda tools. As an example, when minister Jean-Yves Duclos announced the federal contribution to Quebec city's tramway, the Facebook livestream diffusion of the event was plagued by fake likes and comments[146]. Another common phenomenon is "jacking", where one is liking on Facebook a

[142] Statistics Canada (Oct. 25th, 2017), "Census in Brief – Ethnic and cultural origins of Canadians: Portrait of a rich heritage", online: https://www12.statcan.gc.ca/census-recensement/2016/as-sa/98-200-x/2016016/98-200-x2016016-eng.cfm

[143] Lamoureux, Mack (Sept. 11th, 2019), "A QAnon YouTuber Is Running for Office in Canada", *VICE Canada*, online: https://www.vice.com/en_ca/article/8xwxpv/a-qanon-youtuber-is-running-for-office-in-canada

[144] Lamoureux, Mack (Jul. 23rd, 2019), "Far-Right Group Tries to Run for Office, Discovers That Means Outing Themselves", *VICE Canada*, online: https://www.vice.com/en_ca/article/8xzzma/far-right-canadian-nationalist-party-led-by-travis-patron-faces-doxxing-threat

[145] Lamoureux, Mack, "Canadian Far Right Extremism topic", *VICE Canada*, online: https://www.vice.com/en_ca/topic/canadian-far-right-wing-extremism

[146] Clique du Plateau, (Aug. 20th, 2019), "Exclusif: Est-ce que Jean-Yves Duclos a acheté des likes Facebook?", online: https://www.cliqueduplateau.com/2019/08/20/exclusif-est-ce-que-jean-yves-duclos-a-achete-des-likes-facebook/

page, unbeknown to themselves, simply by clicking links on a third-party website, as CBC reporter Reg Sherren learned[147] when a friend asked him why he liked the Conservative party page online.

Besides political parties, many hyper-partisan groups[148] are more than willing to engage in wedge politics online. During the last Ontario provincial elections, the group Ontario Proud worked hard to help Doug Ford's Conservative party defeats then-premier Kathleen Wynne and her Liberal party. Another group, founded by people connected to the founders of Ontario Proud, did the same in the last Quebec provincial elections. In that case, the group Québec Fier, which was rooting for François Legault's Coalition Avenir Québec, used their online and offline platforms to leverage a major campaign against then premier-Philippe Couillard's Liberal Party of Quebec. We cannot attribute the electoral victories of Doug Ford and François Legault only to the work of these hyper-partisan fringe groups. Nonetheless, these groups did play an active role in campaigning during the provincial elections and are now conducting similar efforts during the lead-up to the federal election as well.

Another hyper-partisan group, the online media project *The Rebel Media*, is also playing the game of wedge politics through the circulation of misleading information and has resorted to a number of conspiracy theories[149] such as climate skepticism[150]. While this group usually does not endorse a specific political party, it usually targets and condemns moderate points of view. More notably, if The Rebel is mainly an English-speaking outlet that is read only by a minority, its contributors come from both English- and French-speaking Canada and some of them contribute to other more mainstream medias outlets, especially talk radio programs. The Rebel is the perfect example that shows how hyper-partisan groups can help feed more radical ideas into the general media ecosystem.

Conclusion

All in all, the Canadian elections will be subject to propaganda and disinformation. Some of these misleading sources of information will come from Russia and China, usually regarding specific issues. Ultimately, these issues are unlikely to make or break the election, although they do have the potential to possibly help partially undermine the public's faith in the legitimacy of Canada's democratic institutions. The most important threat to reasoned political debate will come from domestic sources, namely through the advancement of wedge politics by hyper-partisan groups and by politicians themselves. While the Canadian security forces can act against foreign interference, it is the Canadian citizens that will have the burden – and the responsibility – to fight against wedge politics.

---

147 Sherren, Reg (Sept. 15th, 2015), "How I ended up 'liking' the Conservative Party on Facebook without knowing it", *CBC News*, online: https://www.cbc.ca/news/politics/canada-election-2015-like-jacking-facebook-1.3229622

148 Rogers, Kaleigh (Sept. 12th, 2019), "Political disinformation is rampant online. How can voters cope?", *CBC News*, online: https://www.cbc.ca/news/technology/disinformation-political-spin-online-election-2019-1.5279919

149 Warnica, Richard, "Inside Rebel Media", *The National Post*, consulted online in Sept. 2019: https://nationalpost.com/features/inside-ezra-levants-rebel-media

150 Menzies, David (July, 21st, 2019), "The Menzoid's climate change theories make as much sense as the "real" ones", *The Rebel Media*, online: https://www.therebel.media/climate-change-fake-theories-same-effect-real-global-warming-tax-environmentalism

Who Will Bell the Cat?: Interoperability to Combat Digital Threats

Ryan Atkinson, *PhD candidate, Department of Political Science, University of Western Ontario*

The security of digital democracies requires the development of advanced methods for securing traditional institutions, especially as emerging technologies transform all aspects of society. Massive amounts of data are produced about our lives on a daily basis and this increasing stream of information creates vulnerabilities that various actors can take advantage of to gain an unprecedented understanding about individuals and society at large.[151] Emerging technologies have found their way to authoritarian regimes where they are not used to connect and promote the inclusivity of citizens, but to surveil, repress, and curtail dissidents.[152] This danger marks the evermore pertinent need for digital democracies to lead efforts towards developing "a competitive democratic model of digital governance with a code of conduct" that includes "public awareness around information manipulation" to fund "educational programs to build digital critical thinking skills among youth."[153] The present approach focuses on grassroots efforts that have organized to combat the threats of disinformation online, which can be applied to combat other manipulative efforts that take advantage of new avenues of influence made possible by the ubiquitous proliferation of data.

The rampant spread of disinformation on various new media platforms has forced governments to establish disciplinary measures targeting the entities through which such abuses are executed. In some cases, such perpetration is enabled by the designs of the platforms themselves. For example, Facebook's advertisement business model grants advertisers the ability to use sophisticated algorithms to target individuals with specific ads, creating the possibility for political and socio-economic manipulation in targeting people with political messages based on their perceived personalities.[154] Such opportunities have been to the advantage of companies like Cambridge Analytica, a data mining and political consultancy firm that worked on the 2016 Trump Campaign, and gained access to 50 million Facebook profiles to target users with political advertisements.[155]

[151] Bernard Marr, "How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read," *Forbes*, 21 May 2018, https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/#5769f87560ba.

[152] Marcus Michaelsen and Marlies Glasius, "Authoritarian Practices in the Digital Age," *International Journal of Communication,* 12 (2018): 3788-3794, https://ijoc.org/index.php/ijoc/article/viewFile/8536/2458.

[153] Alina Polyakova and Chris Meserole, "Exporting digital authoritarianism: The Russian and Chinese models," *Democracy and Disorder: Foreign Policy at Brookings,* August 2019, https://www.brookings.edu/wp-content/uploads/2019/08/FP_20190827_digital_authoritarianism_polyakova_meserole.pdf.

[154] Avijit Ghosh, Giridhari Venkatadri, Alan Mislove, "Analyzing Political Advertisers' Use of Facebook's Targeting Features," 2019, https://www.ieee-security.org/TC/SPW2019/ConPro/papers/ghosh-conpro19.pdf.

[155] Robinson Meyer, "The Cambridge Analytica Scandal, in Three Paragraphs," *The Atlantic*, 20

Government responses to such scandals have resulted in various sanctions, such as the significant $5 billion fine against Facebook.[156] Nonetheless, many argue that current government efforts against online disinformation, manipulation, and data abuse do not go nearly far enough.[157] Combating this problem requires developing interoperable solutions that must be the central focus to governments, militaries, civil society, and the private sector. This necessity is especially pertinent as the sheer number of disinformation campaigns has doubled in the last two years.[158]

A recent report focusing on government initiatives to counter disinformation globally found that governments are "increasingly aware of the threats posed by disinformation but are struggling to find effective ways to curb its spread."[159] One means of overcoming this struggle is to look to grassroots movements developed by the efforts of independent citizens who, when directly faced with these problems, have fought back with measures ranging from the organizing of volunteers to the development of organizations dedicated to implementing such countering mechanisms. A major value of widespread access to the Internet and additional information communication technologies is the democratization of information, which allows whole investigations to take place utilizing open source information by dedicated teams of volunteers. Examples of such efforts have targeted the dangers of disinformation online to form what can be considered a countermovement, where distinct organizations have operated to counter the threat of disinformation and engage in rigorous fact-checking and evidence-based reporting.

One such organization that has worked tirelessly to combat the threat of online disinformation is Bellingcat. The case will be made for the value of this organization to demonstrate how it can function as a generalizable model for developing further counter responses to other dangers threatening democracies in the Digital Age. Whether these dangers involve the threat that disinformation poses to how individuals understand the world, or to the harvesting of personal data for nefarious ends via political microtargeting. Bellingcat is an investigative journalist organization founded by Eliot Higgins in July 2014 which specializes in fact-checking and Open Source Intelligence (OSINT). It publishes investigative reports online along with case studies and guides on the various techniques used. The name "Bellingcat" is derived from a fable involving a group of mice that are threatened by a cat and decide to hook a bell around its neck to hear it coming. Despite wide support for the initiative, none of the mice are willing to take action to complete the task. None will bell the cat.[160]

One of Bellingcat's first major investigations was into the downing of Malaysian Airlines Flight 17 (MH17). The passenger airline from Amsterdam to Kuala Lumpur was shot down

---

March 2018, https://www.theatlantic.com/technology/archive/2018/03/the-cambridge-analytica-scandal-in-three-paragraphs/556046/.

[156] Rob Davies and Dominic Rushe, "Facebook to pay $5bn fine as regulator settles Cambridge Analytica complaint," *The Guardian*, 24 July 2019, https://www.theguardian.com/technology/2019/jul/24/facebook-to-pay-5bn-fine-as-regulator-files-cambridge-analytica-complaint.

[157] Chris Meserole and Alina Polyakova, "Disinformation Wars," *Foreign Policy*, 25 May 2018, https://foreignpolicy.com/2018/05/25/disinformation-wars/.

[158] Samantha Bradshaw and Philip N. Howard, "The Global Disinformation Order: 2019 Global Inventory of Organized Social Media Manipulation, *Computational Propaganda Research Project, 2019, 2.* https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/09/CyberTroop-Report19.pdf.

[159] Olga Robinson, Alistair Coleman, and Shayan Sardarizadeh, "A Report of Anti-Disinformation Initiatives," *Oxford Internet Institute,* August 2019, https://oxtec.oii.ox.ac.uk/publication/bbc-monitoring-report/.

[160] Aesop's Fables, "Belling the Cat," *Library of Congress,* http://www.read.gov/aesop/003.html.

mid-flight over Eastern Ukraine on 17 July 2014, killing all 298 people onboard.[161] Bellingcat conducted its investigation mainly with volunteers, lacking external funding, and was involved in the discovery of key information which included tracking the responsible BUK missile launcher from Russia to Ukraine and back, locating the field from which the missile was launched, and identifying the separatists involved in Ukraine.[162] Bellingcat published a report with evidence identifying that the BUK was from the 53rd Anti-Aircraft Missile Brigade convoy of the Russian Ground Forces as it moved from Kursk on June 23rd to Millerovo on June 25th.[163] This finding was later echoed by members of the Dutch-led Joint Investigations Team (JIT).[164]

The Bellingcat Investigation team used various videos and photographs for their investigation which had been posted on social media sites that included VKontakte, YouTube, Instagram, and Odnoklassniki. It was determined that the separatists transported the BUK through their controlled territory on July 17th from Donetsk to Snizhne and unloaded it "approximately three hours before the downing of MH17 and was later filmed minus one missile driving through separatist-controlled Luhansk."[165] The convoy was then tracked through video footage in Fedoseyevka, Stary Oskol, Aexeevka, and Olkhovatka on July 19th and 20th.[166] By investigating various photos, videos, and social media posts establishing "time frames, logical routes, and photographic evidence," the Investigation Team was able to track the BUK as it traveled first with the convoy, then into separatists territory, and finally out of Ukraine.[167] Bellingcat demonstrates how direct action can be used to counter the dangers of disinformation, while also using open source intelligence and geolocation to determine the facts of a conflict on the ground. More generally, this kind of independently organized action can be used as a model for additional groups to combat other dangers threatening digital democracies, beyond the scope of on-the-ground conflict fact checking and anti-disinformation initiatives outlined.

Returning to the Cambridge Analytica scandal, the central issue was that inappropriately obtained data was used to study Facebook users' personalities to target political advertising towards them through psychographic profiling.[168] This access was obtained based on an app called "This Is Your Digital Life" created by Cambridge University researcher Aleksandr Kogan in 2014.[169] At the time, Facebook allowed developers to access the profiles of individuals and their friends, enabling Kogan and Cambridge Analytica to

[161] Luke Harding, "Three Russians and one Ukrainian to face MH17 murder charges," *The Guardian*, 19 June 2019, https://www.theguardian.com/world/2019/jun/19/mh17-criminal-charges-ukraine-russia.

[162] Bellingcat Investigation Team, "A Birdie is Flying Towards You: Identifying the Separatists Linked to the Downing of MH17," *Bellingcat,* 2018, https://www.bellingcat.com/wp-content/uploads/2019/06/a-birdie-is-flying-towards-you.pdf.

[163] Bellingcat Investigations Team, "Origin of the Separatists' BUK: A Bellingcat Investigation," 8 November 2014, *Bellingcat,* https://www.bellingcat.com/news/uk-and-europe/2014/11/08/origin-of-the-separatists-buk-a-bellingcat-investigation/.

[164] Mike Corder, "Probe: Missile that downed MH17 came from Russian-backed unit, *Associated Press,* 24 May 2018, https://www.apnews.com/df74081e61374108bba11f39ed504fb9.

[165] Investigations Team, "Separatists' BUK," *Bellingcat*, 8 November 2014.

[166] Veli-Pekka Kivikmaki, "Geolocated July BUK convoy videos in Russia, *Bellingcat*, 7 November 2014, https://www.bellingcat.com/news/uk-and-europe/2014/11/07/geolocated-july-buk-convoy-videos-in-russia/.

[167] Investigation Team, "Separatists' BUK," *Bellingcat*, 8 November 2014.

[168] Sue Halpern, "Cambridge Analytica and the Perils of Psychographics," *The New Yorker,* 30 March 2018, https://www.newyorker.com/news/news-desk/cambridge-analytica-and-the-perils-of-psychographics.

[169] Meyer, "Cambridge Analytica," *Atlantic,* 20 March 2018.

have access to such a large number of user profiles. In the wake of the scandal, some data firms have come together to prevent future such occurrences. Organized by Georgetown University's Institute of Political and Public Service, the effort involves both Republican and Democratic data firms, including: DSPolitical, Bully Pulpit Interactive, NGP VAN, Targeted Victory, DeepRoot Analytics, and WPA Intelligence, among others.[170] Current activities include "promulgating a couple of guiding regulations among data privacy proponents and firms […] to determine what the participants are in favor of […] grinding on a measure that will ensure some clarity for consumers as well as inform them regarding the purpose of their collected information."[171]

The Cambridge Analytica scandal demonstrates a distinct problem facing digital democracies that involves the misuse of media for political ends, compared to the use of opensource methods combating disinformation used by Bellingcat. New movements must develop to combine the operations of related parties to combat such threats. The Georgetown University example demonstrates this with academia and data firms collaborating, not unlike Bellingcat independently investigating MH17 and coming to conclusions that the JIT would later confirm. Interoperability is especially pertinent as similar agents are preparing for upcoming elections with former Cambridge Analytica employees operating new firms, like Data Propria, which are in the employment of the Trump 2020 reelection campaign.[172]

[170] Issie Lapowsky, "Data firms team up to prevent the next CA scandal," *Wired*, Sept 17, 2018, https://www.wired.com/story/political-data-firms-prevent-next-cambridge-analytica/.

[171] "Data companies collaborate to avoid another breach controversy," *Cryptocurry,* https://cryptocurry.com/news/data-companies-collaborate-to-avoid-another-data-breach-controversy/.

[172] Jeff Horwitz, "Trump 2020 working with ex-Cambridge Analytica Staffers," *Associated Press*, 15 June 2018, https://apnews.com/96928216bdc341ada659447973a688e4.

Lights, Cameras, ATMs: Russia's Sandworm and their Contributions to Information Operations

Ian Litschko, *Russia-focused Intelligence Analyst*
Josh Campbell, *Lead Intelligence Analyst, formerly Canadian Armed Forces*

That information has replaced conventional physical might, and wealth, as a dominant form of power is a statement that few would find contentious in our information technology-driven society of constantly-connected citizens. Such an assertion would also seem to be supported through a growing focus by governments, the primary brokers of power, globally to target adversarial information assets.[173] Indeed, even the military jargon for such operations places information in a position of primacy by referring to them as 'information operations,' colloquially 'info ops,' or simply 'IO'. The academic discussion around 'IO' continues to flourish, with many debating its constituent parts,[174] however the practical use of 'IO' is both ongoing and frequent, especially within cyberspace. Perhaps one of the best real-world examples can be found in the active hostility between the Russian Federation, Ukraine, and the broader global defence community. This conflict within Ukraine has seen wide use of 'IO' by the Russians, allegedly disrupting Ukrainian domestic policy,[175] civilian communications,[176] and perhaps most importantly, critical infrastructure.[177] However the objectives for such disruption remain debated.

Russia's targeting of Ukrainian critical infrastructure, on its surface, appears to be a simple flexing of military might, albeit from a cyber paradigm. However, upon closer examination, such targeting reveals multiple goals of the Russian security forces: the application of pressure in support of broader political ends by the Russians against an emerging Ukrainian energy policy; the instillation of fear, uncertainty, and doubt in the population about the capability of Ukraine's ability to defend its infrastructure and its citizens; as well as the disruption of international economic activity within Ukraine. The objective of this paper will be to serve as a starting point to examine the Russian activity targeting Ukraine's critical infrastructure, and the possible motivations and outcomes such targeting would have. However, prior to examining the impact of such efforts, it is important to define the actors who are allegedly responsible: this organization, likely operating as a component of

[173] Kannan, Vishnu. "What Really Happened in the Cyber Command Action Against Iran?" Lawfare. Brookings, July 15, 2019. https://www.lawfareblog.com/what-really-happened-cyber-command-action-against-iran.

[174] Wieck, Brian D., Michael D. Holloway, and Thomas D. Lorenzen. "Information Operations Countermeasures to Anti-Access/Area Denial." The Strategy Bridge, May 11, 2017. https://thestrategybridge.org/the-bridge/?category=#WhatIsIO.

[175] Zinets, Natalia, and Pavel Polityuk. "Ukraine Security Service Accuses Russia of Meddling in Election." Reuters. Thomson Reuters, February 21, 2019. https://www.reuters.com/article/us-ukraine-election-russia/ukraine-security-service-accuses-russia-of-meddling-in-election-idUSKCN1QA1OW.

[176] Polityuk, Pavel, and Jim Finkle. "Ukraine Says Communications Hit, MPs Phones Blocked." Reuters. Thomson Reuters, March 4, 2014. https://www.reuters.com/article/us-ukraine-crisis-cybersecurity/ukraine-says-communications-hit-mps-phones-blocked-idUSBREA231R220140304.

[177] "Critical Infrastructure." Critical Infrastructure | Emergency Management Ontario. Ontario Ministry of the Solicitor General, April 19, 2017. https://www.emergencymanagementontario.ca/english/emcommunity/ProvincialPrograms/ci/ci.html

Russian military intelligence, has been given the *nom de guerre* of 'Sandworm' (amongst many others), and has been responsible for a number of highly sophisticated attacks targeting objectives of national interest to Russia.

It has been convention within the intelligence community, and especially those within the cyber field, to categorize activities that appear to share similarities in tactics, techniques, and procedures (TTP), as well as geographic origins, under various pseudonyms.[178] While such practices were once the exclusive domain of government, this practice has recently been taken up by the private sector both as a marketing tool to name the *bêtes noire* of the cybersecurity industry, as well as a means of tracking groups and actors. However, as no convention exists across the industry, many actors have accrued a myriad of names, including Sandworm, also known as: Telebots, Voodoo Bear, ELECTRUM, Black Energy, and others.

Sandworm is a group that has been active since at least 2013 and has targeted entities with a strategic interest to the Russian Federation. Almost from the outset, Sandworm has differentiated itself as an actor by not only targeting conventional IT assets, but also the so-called industrial control systems (ICS) which are designed to control operations in factories, infrastructure, and building facilities. Two of the most infamous of these alleged ICS operations resulted in the blackouts throughout Ukraine in 2015 and again in 2016. In addition to these, Sandworm has also been alleged to have carried out two of the most disruptive cyber attacks within Ukraine in 2017.

It is because of the high degree of sophistication, perceived motivations, targeting preferences, and TTPs that Sandworm has generally been attributed to the Russian security services, where it is believed that Sandworm falls within the Ministry of Defence's Main Directorate (GU).[179] If accurate, this would be the second cyber actor that has been attributed to the GU, alongside the infamous group, APT28, accused of interfering in the United States' presidential elections of 2016. While fighting under the same master, the differences in motives, targeting preferences, and TTPs each group demonstrates, paints a stark contrast between the groups, ultimately belying a depth of capability that even organizations constrained by the rigors of a bureaucracy are able to achieve.[180] This depth of capability would come to be leveraged against Ukraine, while sending a strong geo-political message regarding Russia's sphere of influence, on the evening of 23 December 2015.

On 23 December 2015, the regional power authority of Ivano-Frankivsk *Oblast PrykarpattyaOblenergo*, was forced to suspend distribution of power affecting nearly 700,000 people. Smaller disruptions were also reported by *Chernivtsioblenergo* (*Chernivtsi Oblast*) and *KyivOblenergo* (Kyiv).[181] [182] This disruption in Ivano-Frankivsk lasted for several hours as a result of a coordinated and targeted cyber attack against these *oblenergos'* infrastructure. Such an attack, not only

[178] "BYZANTINE HADES: An Evolution of Collection." Snowden Doc Search, January 17, 2015. https://search.edwardsnowden.com/docs/BYZANTINEHADESAnEvolutionofCollection2015-01-17_nsadocs_snowden_doc.

[179] Harding, Luke. "How Russian Spies Bungled Cyber-Attack on Weapons Watchdog." The Guardian. Guardian News and Media, October 4, 2018. https://www.theguardian.com/world/2018/oct/04/how-russian-spies-bungled-cyber-attack-on-weapons-watchdog.

[180] Palmer, Danny. "Cyber-Espionage Warning: Russian Hacking Groups Step up Attacks Ahead of European Elections." ZDNet. ZDNet, March 21, 2019. https://www.zdnet.com/article/cyber-espionage-warning-russian-hacking-groups-step-up-attacks-ahead-of-european-elections/.

[181] "Міненерговугілля Має Намір Утворити Групу За Участю Представників Усіх Енергетичних Компаній, Що Входять До Сфери Управління Міністерства, Для Вивчення Можливостей Щодо Запобігання Несанкціонованому Втручанню в Роботу Енергомереж." Ministry of Energy and Coal Industry Ukraine. Ministry of Energy and Coal Industry Ukraine, February 12, 2016. http://mpe.kmu.gov.ua/minugol/control/uk/publish/article?art_id=245086886&cat_id=35109.

[182] Groll, Elias. "Did Russia Knock Out Ukraine's Power Grid?" Foreign Policy, January 8, 2016. https://foreignpolicy.com/2016/01/08/did-russia-knock-out-ukraines-power-grid/.

targeting power generation and distribution, but leveraging cyber as a medium to accomplish it, was an unprecedented action until it occurred again almost one year later. Shortly before midnight, on 18 December 2016, a cyber attack, this time far more complex, was launched, targeting and affecting *Ukrenergo*[183] and leaving nearly a fifth of Kyiv's 2.884 million people without power. The attack targeted the *Pivnichna* substation belonging to *Ukrenergo* and resulted in the station going offline. While initial reports were not able to attribute the attack to the Russian Federation generally, or Sandworm, specifically, the timing and targeting seemed uncanny. Additionally, as more information came to light, several cyber security firms would confirm the suspicions of many: that the second attack in nearly a year was likely the result of the group Sandworm.[184] However, this targeting of ICS and disruption of power was only the initial volley in Sandworm's concerted efforts targeting Ukraine.

On 27 June 2017 dozens of firms around the world reported they had fallen victim to an unknown malicious software (malware) attack that was spreading rapidly throughout their networks.[185] [186] The attack, which was later dubbed *NotPetya,* after the strain of malware leveraged, is estimated to have cost organizations, globally, about $1.2 billion dollars.[187] The attack, it was later discovered, allegedly began with the strategic supply chain compromise, by Sandworm, of the Ukrainian accounting software firm Intellect Service which produced the tax and accounting software called *MeDoc*, and which was widely used in Ukraine or by organizations with Ukrainian employees. The compromise allowed Sandworm to upload their malware into organizations' systems globally, bypassing many of the cyber security controls, and enabling rapid spread within firms' networks. Some of the affected organizations were multinational enterprises, however many of them were small-to-medium sized businesses, including consumer financial organizations, retail outlets, and grocery stores.[188] [189] While the malware itself claimed that functionality of the systems would be restored once a ransom had been paid -- a tactic which is termed *ransomware* -- the reality was that payment of the ransom did not unlock the systems, and only resulted in further financial burdens. The use of ransomware is not new, indeed it has been one of the most prolific types of malware in use between 2014-2017.[190] However, that victims were largely Ukrainian, and that the actors leveraged a

[183] Polityuk, Pavel, Vukmanovic, Oleg, and Jewkes, Stephen, "Ukraine's power outage was a cyber attack: Ukrenergo," Reuters, January 18, 2017, https://www.reuters.com/article/us-ukraine-cyber-attack-energy/ukraines-power-outage-was-a-cyber-attack-ukrenergo-idUSKBN1521BA.

[184] Greenberg, Andy. "Crash Override Malware Took Down Ukraine's Power Grid Last December." Wired. Conde Nast, June 13, 2017. https://www.wired.com/story/crash-override-malware/.

[185] "Case Study: A.P. Møller-Maersk and NotPetya." clearsecuritycomm. Security Communication | Washington | Clear Security Communication, May 23, 2018. https://www.clairetills.com/single-post/2018/05/20/Case-Study-AP-Møller-Maersk-and-NotPetya.

[186] Forrest, Conner. "NotPetya Ransomware Outbreak Cost Merck More than $300M per Quarter." TechRepublic. TechRepublic, October 30, 2017. https://www.techrepublic.com/article/notpetya-ransomware-outbreak-cost-merck-more-than-300m-per-quarter/.

[187] O'Connor, Fred. "NotPetya Still Roils Company's Finances, Costing Organizations $1.2 Billion in Revenue." Cybereason, November 9, 2017. https://www.cybereason.com/blog/notpetya-costs-companies-1.2-billion-in-revenue.

[188] Barysevich, Andrei. "Ukrainian Grocery Store Attacked by #Petya Ransomware Https://T.co/6kbHJ0vZ5G." Twitter. Twitter, June 27, 2017. https://twitter.com/DeepSpaceEye/status/879731426677186562?s=20.

[189] Whitwam, Ryan. "'NotPetya' Ransomware Locking Down Computers Across the World." ExtremeTech, June 27, 2017. https://www.extremetech.com/internet/251711-notpetya-ransomware-locking-computers-across-world.

[190] Cuthbertson, Anthony. "Ransomware Attacks Have Risen 250 Percent in 2017, Hitting the U.S. Hardest." Newsweek. Newsweek, May 28, 2017. https://www.newsweek.com/ransomware-attacks-rise-250-2017-us-wannacry-614034.

sophisticated, targeted and strategic supply chain attack, coupled with the rapidity at which the malware spread, demonstrated that this attack had the hallmarks of Sandworm and that its scale was unprecedented. However, the *NotPetya* attack was not the only Sandworm-led efforts that would target Ukraine in 2017.

On 24 October 2017, only months after the infamous *NotPetya* attack, news of yet another cyber attack began to surface in Europe. This malware, labelled *BadRabbit*, was another ransomware attack which appeared to target primarily Ukraine and selected Russian organizations,[191] as well as affecting companies in Germany, Turkey, Bulgaria, and Japan.[192] The two attacks were not identical, *BadRabbit* was being spread through a fraudulent update to a popular component of web browsers, resulting in far broader impacts versus the more strategically targeted *NotPetya* attack. Additionally, further analysis revealed that unlike *NotPetya*, *BadRabbit* would allow for the recovery of files after payment of the ransom. However, there were a few hallmarks which gave rise to some attributing this to the same actors as *NotPetya*: the code used between the two pieces of malware were *very* similar, and the primary targeting of Ukrainain assets also matched the targeting preferences of Sandworm.[193] [194] While the attacks against critical infrastructure in 2015 and 2016, as well as the broader attacks of *NotPetya* and *BadRabbit* in 2017, have largely been attributed to Sandworm, the objectives of these attacks have remained somewhat opaque, especially to the cyber security community.

While one may declare that the attacks of 2015 and 2016 were initially successful, one may also be tempted to declare the overall efforts of Sandworm, ultimately, wasted: both attacks resulted in the power stations resuming normal operations in hours with seemingly no long-lasting ill-effects and the vulnerabilities exploited by Sandworm, patched. However, attacks of this nature, especially when taken in the broader context of an information operation against a population, are often not merely about the immediate results, but the longer lasting effects such attacks precipitate.

Perhaps the most overt example of these longer lasting effects can be seen in the media outlets' baseless predictions the following year, with many peddling premonitions of a possible attack. Such suspicions belied an unease both in Ukraine and in the broader global community.[195] [196] It is also unlikely that these efforts to target Ukrainian power systems coincided by coincidence with the continued efforts by Russia to prevent Ukraine from diversifying its energy sources in the European Union while slowly weaning itself from

[191] Hern, Alex. "Bad Rabbit: Game of Thrones-Referencing Ransomware Hits Europe." The Guardian. Guardian News and Media, October 25, 2017. https://www.theguardian.com/technology/2017/oct/25/bad-rabbit-game-of-thrones-ransomware-europe-notpetya-bitcoin-decryption-key.

[192] Kessem, Limor, Limor Kessem, and Limor Kessem. "Bad Rabbit Ransomware Outbreak Highlights Risk of Propagating Malware Disasters." Security Intelligence, November 2, 2017. https://securityintelligence.com/bad-rabbit-ransomware-attacks-highlight-risk-of-propagating-malware-outbreaks/.

[193] "Reckless Campaign of Cyber Attacks by Russian Military Intelligence Service Exposed." ncsc.gov.uk. National Cyber Security Centre, October 3, 2018. https://www.ncsc.gov.uk/news/reckless-campaign-cyber-attacks-russian-military-intelligence-service-exposed.

[194] Cimpanu, Catalin. "Security Firms Say Bad Rabbit Attack Carried Out by NotPetya Group." BleepingComputer. BleepingComputer.com, October 25, 2017. https://www.bleepingcomputer.com/news/security/security-firms-say-bad-rabbit-attack-carried-out-by-notpetya-group/.

[195] Sebenius, Alyza. "Will Ukraine Be Hit by Yet Another Holiday Power-Grid Hack?" The Atlantic. Atlantic Media Company, December 13, 2017. https://www.theatlantic.com/technology/archive/2017/12/ukraine-power-grid-hack/548285/.

[196] Zimmerman, Vera. "It's the Holiday Season Again. Will Ukraine Be Ready for the Next Cyberattack?" Atlantic Council, December 21, 2017. https://www.atlanticcouncil.org/blogs/ukrainealert/it-s-the-holiday-season-again-will-ukraine-be-ready-for-the-next-cyberattack/.

a wholly Russian dependence on power.[197] These attacks promulgated a clear message to Ukraine and its citizens that if they pursued this strategic policy, they would face the cold Eastern European winters alone and likely without heat. On the world stage, however, the attacks against Ukraine's critical infrastructure also signalled to the broader defence community that a new paradigm in cyberwarfare had emerged: the targeting of critical infrastructure, even outside periods of direct hostility, was considered fair game.[198] [199]

The attacks involving *NotPetya* and *BadRabbit* however, may have had more tactical and operational objectives, rather than the strategic objectives of the attacks in 2015/16. Instead, it is possible that these attacks were meant as a small-scale show-of-force against the Ukranian government, demonstrating Russia's ability to mobilize rapidly and destabilize the population through affecting consumer necessities like retail, grocery, and consumer finance. These sorts of activities are often intended to sow uncertainty within the population regarding their government and such tactics would be hallmarks of other Russian actors in future information operations.[200] This uncertainty has manifested itself within the Ukranian population[201] and its cyber defence community, with many government agencies warning of impending attacks -- resulting in a 'boy who cried wolf' scenario -- where the Russian wolf simply fails to materialize.[202] [203] [204] [205]

Another potential consequence of the *NotPetya* and *BadRabbit* attacks may have been to instill uncertainty into the global economy both within Ukraine and externally.[206] [207] Many multinational organizations were impacted by the attack -- especially by *NotPetya* -- and it is an attack that would not have otherwise impacted those firms had they not had operations in Ukraine. As such, it

---

[197] Kononczuk, Wojciech, "A dangerous energy policy: Ukraine, despite war, is making itself energy dependent on Russian oil," Energy Post, September 8, 2017, https://energypost.eu/15647-2/.

[198] "Russian Government Cyber Activity Targeting Energy and Other Critical Infrastructure Sectors: CISA." Russian Government Cyber Activity Targeting Energy and Other Critical Infrastructure Sectors | CISA. CISA, March 15, 2018. https://www.us-cert.gov/ncas/alerts/TA18-074A.

[199] Dearden Home Affairs Correspondent @lizziedearden, Lizzie. "Russian Hackers 'Targeting UK's Energy and Communications Networks'." The Independent. Independent Digital News and Media, November 15, 2017. https://www.independent.co.uk/news/uk/home-news/russia-hacking-uk-bt-media-energy-companies-top-spy-security-schief-a8055371.html.

[200] Ajir, Media, and Bethany Vailliant. "Russian Information Warfare: Implications for Deterrence Theory." Strategic Studies Quarterly 12, no. 3 (2018): 70-89. https://www.jstor.org/stable/26481910.

[201] Connel, Michael, and Volger, Sarah, "Russia's Approach to Cyber Warfare," CNA Analysis and Solutions, March 2017, pg 19,https://www.cna.org/CNA_files/PDF/DOP-2016-U-014231-1Rev.pdf.

[202] Polityuk, Pavel. "Ukraine Is Sounding the Alarm on a Potential Russian Cyberattack That Could Rival 'NotPetya'." Business Insider. Business Insider, June 26, 2018. https://www.businessinsider.com/ukraine-cyberattack-russia-notpetya-2018-6.

[203] Brennan, David. "Russia Is Preparing a Huge Cyberattack, Ukraine Warns." Newsweek. Newsweek, June 27, 2018. https://www.newsweek.com/russia-preparing-huge-cyber-attack-ukraine-warns-997170.

[204] Gallagher, Sean. "Ukraine Detects New Pterodo Backdoor Malware, Warns of Russian Cyberattack." Ars Technica, November 20, 2018. https://arstechnica.com/information-technology/2018/11/ukraine-detects-new-pterado-backdoor-malware-warns-of-russian-cyberattack/.

[205] The Daily Beast. "Ukraine, Cisco Warn Russia Is Preparing for a Cyberattack." The Daily Beast. The Daily Beast Company, May 23, 2018. https://www.thedailybeast.com/ukraine-cyber-firms-warn-russia-is-preparing-for-a-cyber-attack.

[206] Blosfield, Elizabeth. "Cyber Business Interruption Remains Area of Uncertainty for Insurance." Insurance Journal, June 12, 2018. https://www.insurancejournal.com/news/national/2018/06/12/491842.htm.

[207] Castellanos, Sara, and Adam Janofsky. "One Year After NotPetya Cyberattack, Firms Wrestle With Recovery Costs." The Wall Street Journal. Dow Jones & Company, June 27, 2018. https://www.wsj.com/articles/one-year-after-notpetya-companies-still-wrestle-with-financial-impacts-1530095906.

would be unsurprising to find that firms within Ukraine generally, and especially those impacted by the attacks, as well as firms considering operations in Ukraine, may have to carefully weigh the consequences of financial investment in Ukraine from a risk perspective.[208] Such an outcome would likely negatively impact and alienate the Ukrainian economy and enable the Russian government to extend and solidify its sphere of influence within Ukraine further still.

Sandworm's involvement in information operations differs from the sort portrayed in the media by groups such as APT28 during the US presidential election of 2016. They do not focus on the spreading of propaganda and disinformation through social and news media, but rather choose to target countries and their citizens more directly. Sandworm's objectives, instead, are conducted through means such disruption of services like critical infrastructure, widespread impact of citizens' day-to-day lives, and unconstrained disruption of global information technology in order to accomplish its ends of information operations facilitating change at a policy level, or sowing uncertainty within the population. The varied nature of Sandworm's activity highlights the diversity of actions that can be taken within military operations, and more specifically information operations, within this space. Ukraine is but one area in which the myriad of Russian state actors operate, and while activities are inherently tailored to the area of operations, they reflect broader patterns that can help shape an understanding and mitigation of future information operations both abroad, and closer to home.

[208] Castellanos, Sara, and Adam Janofsky. "One Year After NotPetya Cyberattack, Firms Wrestle With Recovery Costs." The Wall Street Journal. Dow Jones & Company, June 27, 2018. https://www.wsj.com/articles/one-year-after-notpetya-companies-still-wrestle-with-financial-impacts-1530095906.

The mission of NATO Association of Canada is to promote peace, prosperity, and security through knowledge and understanding of the importance of NATO.

The NAOC has strong ties with the Government of Canada including Global Affairs Canada and the Department of National Defence. We are constantly working to create and maintain relationships with international organizations such as the World Bank Group, the European Bank of Reconstruction and Development, NATO Headquarters, the International Criminal Court, and other prominent international NGOs and think tanks

As a leading member of the Atlantic Treaty Association (ATA), the NATO Association of Canada strives to educate and engage Canadians about NATO and NATO's goal of peace, prosperity and security. NATO Association of Canada ensures that we have an informed citizenry able to contribute to discussions about Canada's role on the world stage.

The NATO Association of Canada
Disinformation and Digital Democracies in the 21st Century

NATO Association of Canada

48 Yonge Street, 610

Toronto, Canada

M5E 1G6

(416) 979-1875

info@natoassociation.ca

www.natoassociation.ca