9-16-2022

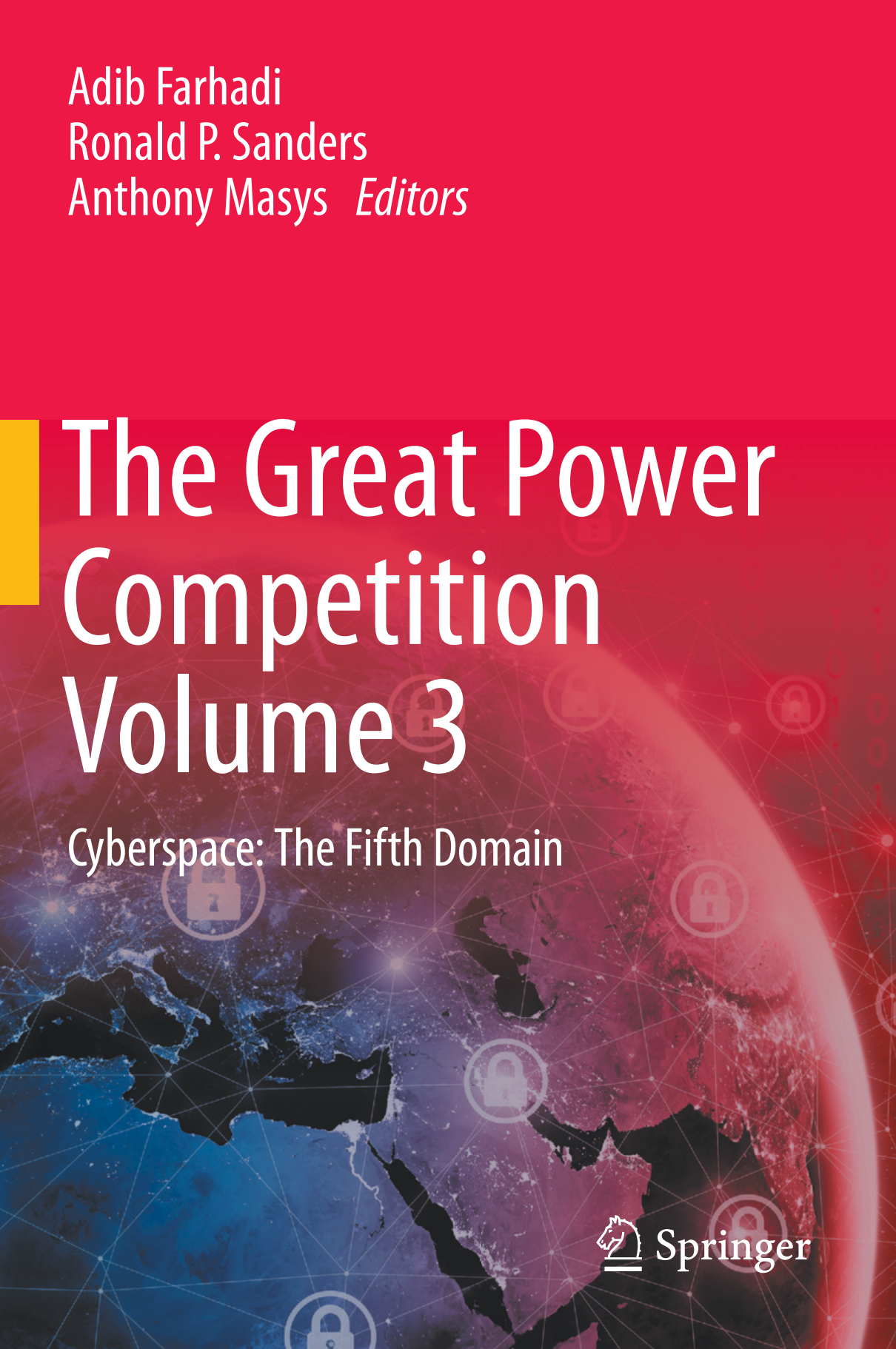
The Future of Cyber-Enabled Influence Operations: Emergent Technologies, Disinformation, and the Destruction of Democracy

Joseph Littell
joseph.littell@westpoint.edu

Adib Farhadi
Ronald P. Sanders
Anthony Masys *Editors*

The Great Power Competition Volume 3

Cyberspace: The Fifth Domain



Springer

The Future of Cyber-Enabled Influence Operations: Emergent Technologies, Disinformation, and the Destruction of Democracy

Joe Littell

Abstract Nation-states have been embracing online influence campaigns through disinformation at breakneck speeds. Countries such as China and Russia have completely revamped their military doctrine to information-first platforms [1, 2] (Mattis, Peter. (2018). China's Three Warfares in Perspective. *War on the Rocks. Special Series: Ministry of Truth.* https://warontherocks.com/2018/01/chinas-three-warfares-perspective/, Cunningham, C. (2020). A Russian Federation Information Warfare Primer. *Then Henry M. Jackson School of International Studies. Washington University.* https://jsis.washington.edu/news/a-russian-federation-information-warfare-primer/.) to compete with the United States and the West. The Chinese principle of "Three Warfares" and Russian Hybrid Warfare have been used and tested across the spectrum of operations ranging from competition to active conflict. With the COVID19 pandemic limiting most means of face-to-face interpersonal communication, many other nations have transitioned to online tools to influence audiences both domestically and abroad [3] (Strick, B. (2020). COVID-19 Disinformation: Attempted Influence in Disguise. *Australian Strategic Policy Institute. International Cyber Policy Center.* https://www.aspi.org.au/report/covid-19-disinformation.) to create favorable environments for their geopolitical goals and national objectives. This chapter focuses on the landscape that allows nations like China and Russia to attack democratic institutions and discourse within the United States, the strategies and tactics employed in these campaigns, and the emergent technologies that will enable these nations to gain an advantage with key populations within their spheres of influence or to create a disadvantage to their competitors within their spheres of influence. Advancements in machine learning through generative adversarial networks [4] (Creswell, A; White, T; Dumoulin, V; Arulkumaran, K; Sengupta, B; Bharath, A. (2017) Generative Adversarial Networks: An Overview. IEE-SPM. April 2017. https://arxiv.org/pdf/1710.07035.pdf.) that create deepfakes [5] (Whittaker, L; Letheren, K; Mulcahy, R. (2021). The Rise of Deepfakes: A Conceptual

J. Littell (✉)
Army Cyber Institute at the West Point, United States Military Academy, West Point, NY 10996, USA
e-mail: joseph.littell@westpoint.edu

A. Farhadi et al. (eds.), *The Great Power Competition Volume 3*,
https://doi.org/10.1007/978-3-031-04586-8_10

Framework and Research Agenda for Marketing. https://journals.sagepub.com/doi/abs/10.1177/1839334921999479.) and attention-based transformers [6] (https://arxiv.org/abs/1810.04805.) (Devlin et al., 2018) that create realistic speech patterns and interaction will continue to plague online discussion and information spread, attempting to cause further partisan divisions and decline of U.S. stature on the world stage and democracy as a whole.

Keywords Disinformation · Cyberspace · Cyber-enabled influence · Influence operation

Executive Summary

Online influence campaigns have become a relatively cheap and impactful way for nations to drive narratives both within their borders and internationally. Since the discovery of Russia and the Internet Research Agency's interference in the 2016 U.S. presidential election and the 2016 U.K. European Union Referendum, seventy other nations have increased their funding for similar campaigns [7].[1] Malign Influence Operations use inauthentic users to push narratives and manipulate social media recommendation algorithms to target key demographics with troves of publicly available data, amplifying already reticent societal fissures and standing as a major danger to democratic institutions throughout the world. Democracy requires four essential functions, as described by Hoover Institute Fellow Larry Diamond. The first requirement is a political system for choosing and replacing the government through free and fair elections. The second requirement is the active participation of people, as citizens, in politics and civic life. The third is the protection of the human rights of all citizens. Finally, the fourth requirement is the rule of law, in which laws and procedures apply equally to all citizens [8].[2] Emergent technologies built from advances in machine learning, such as deepfakes [5],[3] will compound current challenges by adding additional layers of believability to these inauthentic users [9][4] and their malign influence campaigns, making it more difficult for audiences to discern fact from fiction.

Russia and China have utilized these techniques heavily across the last two decades and expanded on them greatly in the COVID era. Other smaller nations have flocked

[1] Bradshaw, S; Bailey, H; Howard, P. (2020). 2020 Global Inventory of Organized Social Media Manipulation. *Computational Propaganda Research Project. Oxford Internet Institute. Oxford University.* https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/127/2021/02/CyberTroop-Report20-Draft9.pdf.

[2] https://diamond-democracy.stanford.edu/speaking/lectures/what-democracy.

[3] Whittaker, L; Letheren, K; Mulcahy, R. (2021). The Rise of Deepfakes: A Conceptual Framework and Research Agenda for Marketing. https://journals.sagepub.com/doi/abs/10.1177/1839334921999479.

[4] Bastos, M., & Mercea, D. (2018). The public accountability of social platforms: lessons from a study on bots and trolls in the Brexit campaign. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, *376*(2128), 20,180,003. https://doi.org/10.1098/rsta.2018.0003.

to methods to bolster their spheres of influence as well as confound international audiences due to the low cost of entry as well as almost relatively nonexistent regulatory barriers to participate at scale. Numerous prevalent examples exist during the last year. Russian attempts at interfering with the 2020 U.S. Presidential election through planting false stories and pushing various false narratives of a stolen election are well known [10].[5] China has used online influence campaigns to crush popular support for democracy in Hong Kong [11],[6] limiting knowledge and creating confusion around allegations of ethnic cleansing and genocide against the Muslim Uyghurs population in Xinjiang province [12],[7] and revitalize its global image in the wake of the COVID19 pandemic [13].[8] Azerbaijan has effectively used disinformation during its ongoing military excursions into Armenia [14].[9] Myanmar's military coup and consolidation of power removed democratically elected officials while killing ethnic minorities in the north and west of the country [15].[10]

As confusion around these events continues, it undermines a basic understanding of the situation and erodes the ability of governments to act collectively. Groups like Bellingcat use collaborative efforts to teach tactics to investigate and rapidly disseminate their findings to a global audience. The thankless efforts of subject matter experts and analysts will unfortunately not be enough to prevent the fomenting of dissidence against democratic ideals. We need active development of machine learning algorithms to handle the flood of coordinated influence campaigns, education on the effects these campaigns have on our population, and expertise and human input to make any sort of headway. Coordination efforts would require a central agency or government department to bring together the vast and various elements of academia, business, and government efforts already underway so that they may share best practices and efforts amongst each other.

---

[5] National Intelligence Council. (2021). Foreign Threats to the 2020 US Federal Elections. *Intelligence Community Assessment.* https://www.dni.gov/files/ODNI/documents/assessments/ICA-declass-16MAR21.pdf.

[6] Wood, D; McMinn, S; Feng,E. (2019). China Used Twitter to Disrupt Hong Kong Protests, but Efforts Began Years Earlier. *National Public Radio.* https://www.npr.org/2019/09/17/758146019/china-used-twitter-to-disrupt-hong-kong-protests-but-efforts-began-years-earlier.

[7] Uyghur Human Rights Project. (2020). The Happiest Muslims in the World: Disinformation, Propaganda, and the Uyghur Crisis. *Uyghur Human Rights Project*. https://docs.uhrp.org/pdf/Disinformation_Propganda_and_the_Uyghur_Crisis.pdf.

[8] Bernard, R; Bowsher, G; Sullivan, R; Gibson-Fall, F. (2021). Disinformation and Epidemics: Anticipating the Next Phase of Biowarfare. *Health Security. Volume 19, Number 1.* https://www.liebertpub.com/doi/pdf/10.1089/hs.2020.0038.

[9] Giles, C; Bhat, U. (2020). Nagoro-Karabakh: The Armenian-Azeri 'Information Wars'. *BBC Reality Checks and Anti-disinformation Unit.* https://www.bbc.com/news/world-europe-54614392.

[10] Beech, H. (2021). 'Now We Are United': Myanmar's Ethnic Divisions Soften after Coup. *New York Times.* https://www.nytimes.com/2021/04/30/world/asia/myanmar-ethnic-minority-coup.html.

Introduction

A once-in-a-century global pandemic swept across the world, shuttering economies [16],[11] killing millions, and infecting millions more [17][12] with long-term health complications that are, as of yet, not fully understood [18].[13] Entire countries have shut down normal day-to-day habits to quell the spread of a novel coronavirus with well-documented origins from China. Mass transit, eating in a restaurant, watching movies in a theatre, and most typical extended family get-togethers came to a halt to stop the spread of a dangerous respiratory virus [19].[14] Phrases such as "social distancing," "mask-up," and "stop the spread" became a part of common vernacular. As most nations scrambled to keep their citizens safe and control the situation to the best of their ability, China and Russia saw the chaos unfolding as an opportunity to spread disinformation surrounding COVID19.

Russia began its disinformation campaigns to sow disarray among democratic nations both near and far. Initially fueled by underplaying the threat of COVID19 [20],[15] they quickly began going after COVID19 restrictions and masking requirements, exacerbating a burgeoning anti-lockdown movement [21][16] across the United States and parts of Europe [22].[17] Their goal: have Americans and Europeans question their governments' attempts to protect the population as some form of removal of human rights. As the monumental push to create, test, and distribute a vaccine to protect the population from the virus came with unparalleled cooperation between

[11] McKinsey and Company. (2021). Covid-19: Implications for Business. *McKinsey and Company. Executive Briefing.* https://www.mckinsey.com/business-functions/risk-and-resilience/our-insights/covid-19-implications-for-business#.

[12] Center for System Sciences and Engineering (CSSE). (2021). COVID-19 Dashboard. *Johns Hopkins University. ARCGIS.* https://www.arcgis.com/apps/dashboards/bda7594740fd40299423467b48e9ecf6.

[13] Blomberg, B; Mohn, K; Brokstad, KA et al. (2021). Long COIVD in a prospective cohort of home-isolated patients. *Nature Medicine.* https://www.nature.com/articles/s41591-021-01433-3.

[14] Centers for Disease Control and Prevention. (2020). CDC Calls on Americans to wear mask to prevent COIVD-19 Spread. *Centers for Disease Control and Prevention.* https://www.cdc.gov/media/releases/2020/p0714-americans-to-wear-masks.html.

[15] Loucaides, D; Perrone, A. (2021). Germany's COVID sceptics fueled by Russian media and far-Right Conspiracies. *openDemocracy.* https://www.opendemocracy.net/en/germanys-covid-sceptics-fuelled-by-russian-media-and-far-right-conspiracies/.

[16] Benson, T. (2020). Trolls and bots are flooding social media with disinformation encouraging states to end quarantine. *Business Insider. Tech.* https://www.businessinsider.com/trolls-bots-flooding-social-media-with-anti-quarantine-disinformation-2020-4.

[17] Emmot, R. (2020). Russia deploying coronavirus disinformation to sow panic in West, EU document says. *Reuters.* https://www.reuters.com/article/us-health-coronavirus-disinformation/russia-deploying-coronavirus-disinformation-to-sow-panic-in-west-eu-document-says-idUSKBN21518F.

government, academia, and industry in a record time, Russia moved to prop up long-standing anti-vaccination conspiracy theories to slow the adoption of the COVID19 vaccine [23].[18]

China, by contrast, looked less to pull the rest of the world down but rather to elevate its own standing. Downtrodden by the news reports about Chinese failures to initially contain the virus as the source of the pandemic, China began a disinformation campaign in February 2020 with multiple, spontaneous origins to lessen their fault [24].[19] China bolstered the campaign through additional influence operations, sending medical supplies and doctors to Europe [25][20] and later vaccines to South America [26].[21] As the number of U.S. and E.U. cases and deaths increased, China used its social media sphere to antagonize the perceived failures of democracy to handle the virus in comparison to their "superior system of government" [24][22] When these attempts failed, China decided to claim instead that COVID19 was actually a bioweapon from the United States, part of a conspiracy to destroy China [27].[23]

While the two competing powers had different goals, their use of social media and disinformation stood closely aligned. Inauthentic accounts, astroturfing, information laundering, and algorithmic manipulation were mainstays used to push messages further and faster. A new era of propaganda and disinformation is coming of age, ready to divide citizens and discredit democracy. Without properly understanding the problem and how it will be used more readily by nation-states in addition to the traditional powers, it will be near impossible to stop.

---

[18] Miller, H. (2021). Russia's Anti-Vaccine Propaganda is Tatamount to a Declaration of War. *Center for Medical Economic and Innovation.* https://medecon.org/russias-anti-vaccine-propaganda-is-tantamount-to-a-declaration-of-war/.

[19] Chen, E. (2021). Chinese COVID-19 Misinformation A Year Later. *The Jamestown Foundation Global Research and Analysis.* https://jamestown.org/program/chinese-covid-19-misinformation-a-year-later/.

[20] Reuters Staff. (2020). China sends medical supplies, experts to help Italy battle coronavirus. *Reuters Healthcare and Pharma.* https://www.reuters.com/article/us-health-coronavirus-italy-respirators/china-sends-medical-supplies-experts-to-help-italy-battle-coronavirus-idUSKBN2101IM.

[21] Mallapaty, S. (2021). China's COVID vaccines are going global—but questions remain. *Nature Magazine* https://www.nature.com/articles/d41586-021-01146-0.

[22] Chen, E. (2021). Chinese COVID-19 Misinformation A Year Later. *The Jamestown Foundation Global Research and Analysis.* https://jamestown.org/program/chinese-covid-19-misinformation-a-year-later/.

[23] Kinetz, E. (2021). Anatomy of a conspiracy: With COVID, China took leading role. *Associated Press.* https://apnews.com/article/pandemics-beijing-only-on-ap-epidemics-media-122b73e134b780919cc1808f3f6f16e8.

Background

Propaganda and disinformation are not new methods for changing the behavior of a population, nor are they relegated to nonmilitary roles and impacts. In "Munitions of the Mind," author Philip Taylor states that the use of military poems and hymns to build morale and strike fear into adversaries was commonplace within ancient Babylonia by 1300 BCE [28].[24] The Assyrians, espec ially, used these methods heavily during their campaigns against the nation-states of the Kassites, Ur, and Urak to build their armies' confidence while making their enemies fear their great feats on the battlefield [28].[25] A natural progression of the adage, "history is written by the victors," continued until the invention of the Gutenberg printing press sparked a revolution in the ability to spread information.

Although printing methods preceded Gutenberg's landmark invention, 1445 CE stands as a crossroads of propaganda [29].[26] No longer were nations or the church responsible for the spread of information and knowledge; individuals with access to printing presses could furnish their own views en masse to the greater population. Text—and literacy—broke free from the original Latin as scholars, scientists, and religious figures could now produce works in their native tongues [30].[27] It is no wonder that the Renaissance and Reformation, two monumental social upheavals of modern history, took place in the wake of the printing press. Control of information has been a critical part of nations' stories—the only difference now is the medium through which they spread.

Pushback against the Catholic church throughout Europe was led, in part, by propaganda from Protestant believers. The seismic power shift was primarily due to the change in technologies that opened up access to information and who could shape information into shared knowledge. The advent of radios and motion pictures led to their immediate use as media for propaganda. The Nazi Party utilized radio heavily to consolidate power following 1933 [31].[28] Both Axis and Allied powers used newsreels at the start of motion pictures to spread propaganda during World War II [32][29] and used Hollywood to spread American influence thereafter. It, therefore, was the logical outcome that a new mode of communication—the internet, and especially social media—shepherded in a new age of propaganda and disinformation.

---

[24] Taylor, P. (2003). Munitions of the mind: A history of propaganda. *Manchester University Press.* https://www.jstor.org/stable/j.ctt155jd69.

[25] Taylor, P. (2003). Munitions of the mind: A history of propaganda. *Manchester University Press.* https://www.jstor.org/stable/j.ctt155jd69.

[26] Edwards, M. (1994). Printing, Propaganda, and Martin Luther. *Berkeley: University of California Press.* https://publishing.cdlib.org/ucpressebooks/view?docId=ft3q2nb278;chunk.id=0;doc.view=print.

[27] https://publishing.cdlib.org/ucpressebooks/view?docId=ft3q2nb278;chunk.id=0;doc.view=print.

[28] https://www.ushmm.org/collections/bibliography/nazi-propaganda-1.

[29] Jowett, G; O'Donnell, V. (2012). Propaganda and Persuasion, 5th Edition. *Sage Publications.*

When applied to the United States and Western liberal democracies as a whole, propaganda and disinformation can attack the foundational elements and requirements needed for self-governance directly. If a population is led to believe that elections are stolen, that, in turn, leads to a breakdown in participation in the democratic process. If that process is skewed to one individual group or belief structure, it will lead to the dissolution of rights or protections of subsets of the citizenry under the rule of law. This process of breaking down the foundation of democracy is akin to removing the leg of a table: while the structure may remain mostly intact, it cannot hold weight, and its contents fall, slipping into disarray and chaos.

Inauthentic Users and Astroturfing

In a May 2021 presentation to the U.S. Agency for Global Media and the Aspen Institute, Peter Pomerantsev spoke of how, ideally, democracies operate under the notion of one vote, one voice; but with online disinformation campaigns, adversarial groups and countries can drown out an individual's voice with a flood of activity from inauthentic users [33].[30] This process of building a false consensus is commonly referred to as "astroturfing," a play on a grassroots initiative in that some outside force is manufacturing the consensus to make it seem as if it is authentic. An example of astroturfing that was uncovered was around a movement to discredit feminists through the Twitter hashtag #endfathersday [34].[31] Several inauthentic accounts claiming to be women, particularly women of color, began posting vitriolic speech towards men under the hashtag. Members of the platform quickly began calling out inauthentic users under the hashtag #YourSlipIsShowing.

Inauthentic users are those who use the anonymity of the internet to create fake personas to push ideas, disinformation, and state propaganda. These false users can drive the astroturfing initiative to drown out authentic views of a situation, pushing forth a narrative that is most beneficial to the malign influence campaigns' goals. Inauthentic users fall under two broad categories, automated false accounts or "bots," and human-run false accounts or "sock puppets" [7].[32]

Bots are defined as accounts that are created and their actions automated to resemble those actions of a genuine user. These bots can be quickly created and

[30] https://watch.eventive.org/m4df2021/play/6074567c86f143003e0cbd99/6074697b69154f003086bb13.

[31] Broderick, R. (2014). Activists Are Outing Hundreds Of Twitter Users Believed To Be 4chan Trolls Posing As Feminist. *BuzzFeed News*. https://www.buzzfeednews.com/article/ryanhatesthis/your-slip-is-showing-4chan-trolls-operation-lollipop.

[32] Bradshaw, S; Bailey, H; Howard, P. (2020). 2020 Global Inventory of Organized Social Media Manipulation. *Computational Propaganda Research Project. Oxford Internet Institute. Oxford University.* https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/127/2021/02/CyberTroop-Report20-Draft9.pdf.

employed into an information space for both spreading disinformation and manipulating engagement [35],[33] and algorithmic confounding [36].[34] Due to their reliance on automation, much of their interactions becomes easier to track due to temporal aspects, posting at regular intervals, social subgraphs (who they are interacting with), and similarity of content, as it is easier to repeat content than create new, unique content for so many additional users [37].[35] That said, advances in technology, which will be addressed later in the chapter, are making these trackable features more difficult.

Bots then contrast with sock puppets accounts, which are operated by real people under a fake persona. This allows the accounts to have a more authentic and personalized feel. Although tracking by metadata and I.P. addresses are still possible, the variety of the content produced, and ability to converse with anyone interacting with the influence campaign more directly, sock puppet accounts can be more difficult to ascertain.

Although bots, sock puppets, and so-called "troll farms" came into national media attention and prominence following the 2016 U.S. presidential election and British European Union Referendum, their use has a long history. The earliest documented attempts were on Russian Politicheskie web forums in the late 1990s and early 2000s [38].[36] Journalist Anna Polyanskaya described the shift from liberal democratic themes to strict authoritarian ideals in her expose in Vesnik in 2003. Polyanskaya observed a notable shift as Russian President Vladimir Putin began to consolidate his political power, citing various inconsistencies in the overall tone of posts, which leaned pro-Putin, compared to polling across RuNet, which leaned liberal and was constrained by single votes per I.P. address [38].[37]

Later in the 2000s, a Russian arm of the hacker collective known as Anonymous released emails tied to the pro-Kremlin youth organization, Nashi [39].[38] These emails described web "brigades" of paid internet users who flooded various elements of the internet to counter claims and harass dissidents. These users were paid upwards of $3 per post depending on the amount of feedback [39].[39] By the 2010s, Russia

---

[33] Ferrara E. (2020) Bots, Elections, and Social Media: A Brief Overview. In: Shu K., Wang S., Lee D., Liu H. (eds) Disinformation, Misinformation, and Fake News in Social Media. *Lecture Notes in Social Networks. Springer, Cham.* https://doi.org/10.1007/978-3-030-42699-6_6.

[34] Bakir, V. and McStay, A. (2018). Fake News and the Economy of Emotions. *Digital Journalism, 6:2, 154-175,* https://doi.org/10.1080/21670811.2017.1345645.

[35] Anderson, J; Rainie, L; Vogels, e > (2021). Expers Say the 'New Normal' in 2025 Will Be Far More Tech-Driven, Presenting More Big Challenges. Pew Research Center. *Emerging Technologies. Future of the Internet.* https://www.pewresearch.org/internet/2021/02/18/experts-say-the-new-normal-in-2025-will-be-far-more-tech-driven-presenting-more-big-challenges/.

[36] Polysanskaya, A; Krivov, A; Lomko, I. (2003). The Virtual Eye of Big Brother. *Vestnik Online.* http://www.vestnik.com/issues/2003/0430/win/polyanskaya_krivov_lomko.htm.

[37] Polysanskaya, A; Krivov, A; Lomko, I. (2003). The Virtual Eye of Big Brother. *Vestnik Online.* http://www.vestnik.com/issues/2003/0430/win/polyanskaya_krivov_lomko.htm.

[38] Elder, M. (2012). Polishing Putin: hacked emails suggest dirty tricks by Russian youth group. *The Guardian.* https://www.theguardian.com/world/2012/feb/07/putin-hacked-emails-russian-nashi.

[39] Elder, M. (2012). Polishing Putin: hacked emails suggest dirty tricks by Russian youth group. *The Guardian.* https://www.theguardian.com/world/2012/feb/07/putin-hacked-emails-russian-nashi.

moved beyond its internal manipulation and began heavily utilizing online brigades against both regional neighbors and internationally against the U.S. and West as a whole. By 2014, Russia was fomenting large-scale operations in support of hybrid warfare through the support of far-right and ethnocentric movements in the U.S. and Europe.

Although Russia was the first suspected peddler in online manipulation, they were not the only country entering into this burgeoning field. China began utilizing online manipulation through paid internet commentators as early as October 2004, when the CCP Propaganda Department of Changsha began hiring users to manipulate local online forums [40].[40] Quickly, the national CCP Ministry of Education began censoring and astroturfing, or creating fake grassroots movements, across Chinese university bulletin board systems and forums [41].[41]

Networks of inauthentic users working together utilize various tools available to them to spread disinformation and propaganda to susceptible users in order to muddy the waters [42],[42] gain an information advantage [43],[43] or enact a real world response [44].[44] Typically, they coordinate their functions to use social media platforms recommendation algorithms to their advantage [45],[45] to move fringe ideas and disinformation from their inauthentic network to more authentic ones to add legitimacy to the disinformation campaign [46],[46] thus, creating more misinformation as it spreads. These networks can do this with nightmarish precision due to the vast mountains of information available for sale [47],[47] stolen from email servers or

[40] Bandurski, D. (2008). China's Guerrilla War for the Web. *China Media Project. China Internet Research Conference.* https://chinamediaproject.org/2008/07/07/feer-chinas-guerrilla-war-for-the-web/.

[41] King, G; Pan, J; Roberts, M. (2017) > How the Chinese Government Fabricates Social Media Post for Strategic Distraction, not Engaged Argument. *American Political Science Review, vol. 111, pg. 484–501.* https://gking.harvard.edu/files/gking/files/50c.pdf.

[42] Lock, I; Ludolph, R. (2019). Organizational propaganda on the Internet: A systematic review. *Public Relations Inquiry. vol 9, issue 1, pg 103–127.* https://journals.sagepub.com/doi/full/10.1177/2046147X19870844.

[43] Kenny, C; Bergman, M. (2019). Understanding and Combating Russian and Chinese Influence Operations. *Center for American Progress, Foreign Policy and Security.* https://www.americanprogress.org/issues/security/reports/2019/02/28/466669/understanding-combating-russian-chinese-influence-operations/.

[44] Select Committee on Intelligence. United States Senate. (2018). Russian Active Measures Campaigns and Interference in the 20,216 U.S. Election Volume 2: Russia's Use of Social Media with Additional Views. *116*th *Congress. Senate. 1*st *Session. Report 116-XX.* https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume2.pdf.

[45] Meserole, C. (2018). How misinformation spreads on social media—and what to do about it. *Brookings Institute.* https://www.brookings.edu/blog/order-from-chaos/2018/05/09/how-misinformation-spreads-on-social-media-and-what-to-do-about-it/.

[46] Arjomand, N. (2019). Information Laundering and Globalized Media—Part I: The Problem. *Center for International Media Assistance.* https://www.cima.ned.org/blog/information-laundering-and-globalized-media-part-i-the-problem/.

[47] Sherman, J. (2021). Federal Privacy Rules Must Get "Data Broker" Definitions Right. Lawfare. *Consumer Data.* https://www.lawfareblog.com/federal-privacy-rules-must-get-data-broker-definitions-right.

scraped from publicly available sources [48][48] from the very same platforms they are manipulating. By building this false notion of consensus through astroturfing, real voices are drowned out, directly hindering the ability of citizens to play a part in active participation within civil society [49].[49] Without the ability to have a shared foundation between citizens, the ability to engage in the collective action of self-governance faces a significant threat of being destroyed.

Algorithmic Confounding

Search engines [50],[50] e-commerce platforms [51],[51] and social media networks [52][52] all utilize machine learning algorithms to recommend items to their users. These items range from a new book or movie to the news they read. These algorithms are split into three main types: content filtering, collaborative learning, and knowledge-based systems [53].[53] Each algorithm works slightly differently and can be used individually or in tandem with the other base algorithms to tailor the end user's experience. Content filtering, simply put, is using the choices of a single user and finding similarly categorized items to recommend, like offering a keyboard with a purchase of a computer mouse. Collaborative learning, also called collaborative filtering, takes the preference of one user and compares it to other similar users. The users with a large overlap in preference may be driven to other items frequented by like-minded users. For example, if one user enjoys a particular rock band's music, they may be recommended a different band's music because another user also enjoys the first band and regularly listens to the recommended band as well. Finally, knowledge-based systems require rules written by subject matter experts with vast amounts of specific domain knowledge in the product being recommended. Usually, knowledge-based systems are hybrid algorithms that work with content modeling, collaborative learning, or both.

---

[48] The National Law Review. (2019). Data Scraping Survives! (At Least for Now) Key Takeaways from 8th Circuit Ruling on the HIQ vs. LinkedIn Case. *The National Law Review.* https://www.natlawreview.com/article/data-scraping-survives-least-now-key-takeaways-9th-circuit-ruling-hiq-vs-linkedin.

[49] Diamond, L. (2004). What is Democracy. *Lecture to Hilla University for Humanistic Studies.* https://diamond-democracy.stanford.edu/speaking/lectures/what-democracy.

[50] Google Developers. (2021). Introduction to Recommendation Systems. *Google.* https://developers.google.com/machine-learning/recommendation.

[51] Hardesty, L. (2019). The history of Amazon's recommendation algorithm. *Amazon Science.* https://www.amazon.science/the-history-of-amazons-recommendation-algorithm.

[52] Ilic, A. Kabiljo, M. (2015) Recommending items to more than a billion people. *Facebook Engineering, Core Data, ML Applications.* https://engineering.fb.com/2015/06/02/core-data/recommending-items-to-more-than-a-billion-people/.

[53] Oracle. (2020). Introduction to Recommendation Engines. *Oracle Cloud Infrastructure Data Science.* https://www.oracle.com/a/ocom/docs/introduction-to-recommendation-engines-wp.pdf.

These recommendations alone have a profound effect on marketing and the end user's decision-making process. Researchers from Princeton University's departments of computer science and sociology showed that recommendation engines increased homogeneity through the selection bias and reinforced feedback from retraining the algorithm, commonly referred to as live systems, on those initial choices, making them more skewed in their distribution of items recommended [54].[54] When applied to controversial or political topics, it begins to skew the populace from a shared basis of understanding. This can be leveraged to make greater divides between sets of people despite not being as far apart as they might seem. Without a shared perspective and culture, consensus building becomes much more difficult, if not impossible. Without consensus, a representative democracy will begin to burst at the seams.

In Brazil, recommendations from YouTube helped to radicalize numerous individuals after updates to its recommendation engine. New York Times' reporters Max Fisher and Amanda Taub traced how Brazilian far-right politicians and conspiracy theorists came to prominence by embracing the video platform, elevating their exposure through the use of music and video games while exposing political beliefs. This allowed once-fringe lawmakers like Jair Bolsonaro to become president [55][55] and others to win in dramatic landslides. YouTube's algorithm drives upward of 70 percent of viewership, keeping users watching, and driving billions of dollars in ad revenue to parent company Alphabet [56][56] YouTube also acts as a producer of the content, since their ad revenue sharing pays the influencers to create and publish content. As such, it adds liquidity to some of these operations, funding those who otherwise might not be able to push the disinformation.

Engagement on most of the major platforms is driven by some form of user voting system. Users can like, upvote, share, and retweet content to make it more visible in the recommendation system. This can and is easily manipulated to twist public opinion on controversial issues. In 2014, the terrorist group known as the Islamic State of Iraq and Syria, or ISIS as it is more commonly referred, began taking over popular hashtags, a grouping mechanism for similar topics, to espouse their violent ideology and recruit new members [57].[57] Similarly, on Reddit, a group of users and moderators manipulated Reddit's use of stickied posts, or posts that are the first seen on a smaller "subreddit," to create an inorganic flow of engagement. This caused posts on r/The_Donald, a subreddit for then U.S. presidential candidate Donald Trump, to surge to the top of r/all, the most popular posts throughout the platform

---

[54] Chaney, A; Stewart, B; Engelhardt, B. (2018). How Algorithmic Confounding in Recommendation Systems Increases Homogeneity and Decreases Utility. *Twelfth ACM Conference on Recommender Systems (RecSys'18).* https://scholar.princeton.edu/bstewart/publications/how-algorithmic-confounding-recommendation-systems-increases-homogeneity-and.

[55] Fisher, M; Taub, A. (2019). How YouTube Radicalized Brazil. *New York Times.* https://www.nytimes.com/2019/08/11/world/americas/youtube-brazil.html.

[56] Solsman, J. (2018). YouTube's AI is the puppet master over most of what you watch. *CNET*. https://www.cnet.com/news/youtube-ces-2018-neal-mohan/.

[57] Girginova, K. (2017). Hojacking Heads and Hashtags. *Global Journal, vol 10, Issue 56.* https://globalejournal.org/global-e/august-2017/hijacking-heads-hashtags.

starting in March 2016. This continued throughout the summer and the election until November 24 of that year when the CEO of Reddit, Steve Huffman, intervened [58][58] and stopped the manipulation.

Coupling the user manipulation of engagement with the algorithmic confounding present in most recommendation engines allows for disinformation peddlers to move their message from the white noise of social media to virality. While fringe materials may not be naturally seen by the majority of the populace, abusing these systems can vastly increase the number of people who view false and misleading materials. A study by Massachusetts Institute of Technology's Media Lab, authors Soroush Vosoughi, Deb Roy, and Sinan Aral found that false news stories spread further and faster than true stories [59][59]; reaching upwards of 100,000 users at the top 1% of stories, where traditional news at the same level rarely peaked 1,000 users. False stories also spread at a rate 6 times that of their truthful counterparts.

Information Laundering

One of the biggest hurdles our adversaries face is making their disinformation reaches a critical mass. While astroturfing and algorithmic confounding can be used to get more individuals to see the content, information laundering will spread the message more quickly due to the added legitimacy [60].[60] Information laundering is the act of a legitimate source, whether a political commentator, social media influencer, or politicians themselves, taking disinformation and spreading it across their network. By attaching their name or brand to a set of disinformation, this acts to legitimize it even if the user in question never directly states they believe in it. This methodology was shown in the 2021 report from the Center for Countering Digital Hate called the Disinformation Dozen, where 12 individuals were responsible for up to 63% of anti-vaccination content on social media [61].[61]

The use of intermediaries in many cases invokes a parasocial relationship between the trusted influencer, and the target audience [62].[62] The audience trusts the influencer's opinion as they would any other real world interpersonal relationship, causing

[58] Huffman, S. (2016). TIFU by editing some comments and creating an unnecessary controversy. *Reddit Announcements.* https://www.reddit.com/r/announcements/comments/5frg1n/tifu_by_editing_some_comments_and_creating_an/.

[59] Vosoughi, S; Roy, D; Aral, S. (2018). The spread of true and false news online. *Science vol. 359. Issue 6380 pp. 1146–1151.* https://science.sciencemag.org/content/359/6380/1146.

[60] Meleshevich, K. Schafer, B. Online Information Laundering: The Role of Social Media. *Alliance for Securing Democracy. The German Marshall Fund of the United States.* https://securingdemocracy.gmfus.org/online-information-laundering-the-role-of-social-media/.

[61] Center for Countering Digital Hate. (2021). This Disinformation Dozen. *Center for Countering Digital Hate.* https://www.counterhate.com/disinformationdozen.

[62] Parasocial interaction. https://www.oxfordreference.com/view/10.1093/oi/authority.20110803100305809.

strangers to have a much stronger hold on citizen's lives than there would be in traditional outreach and marketing. With the increase of social media usage over time particularly through the 2020 global SARS COVID-19 pandemic [63],[63] more and more people relied on social media to get their much-needed interpersonal relationships. This reliance led to a huge growth in disinformation spread and acceptance of conspiracy theories, particularly surrounding anti-government and anti-democracy movements [64].[64]

Information laundering acts as a pipeline, moving disinformation from an untrusted source to a trusted one for greater dissemination and belief. Bret Schafer and Kirill Meleshevich of the Alliance for Securing Democracy describe information laundering as a three-step process [60].[65] The first step is placement. This is done through the inauthentic accounts, whether bots or sockpuppets, putting the initial disinformation into the information environment. Typically, bots will be used for acting in tandem to spread the disinformation through networks, while simultaneously using some form of algorithmic confounding to elevate it to multiple users. By using multiple accounts in tandem, the disinformation is made legitimate since it becomes more difficult to attribute it to a single source, and therefore debunk.

The next step is layering. In order to gain authenticity, and the potential removal of content, the disinformation must travel through various layers, usually some form of intermediary to spread. In many cases the intermediaries and unwitting participants who spread misinformation because it fits a social narrative that they agree with. While algorithmic confounding can help with this, it is usually not the only pathway. For example, during the lead up in the 2016 U.S. Presidential election, Russian hackers broke into the Democratic National Convention's chairman Jon Podesta's personal and professional emails, and passed selected information to Wikileaks, a seemingly neutral third party [65].[66] Wikileaks' founder, Julian Assange, was heavily tied to Russian intelligence agencies and often acted as a disinformation conduit for them, going as far to obscure the true source of the leaked emails, insinuating it was a former DNC staffer, Seth Rich, who was tragically murdered in July of 2016 [66].[67]

The final step of information laundering is integration. This step requires legitimate news outlets to share the disinformation to a greater scale. In 2020 alone,

---

[63] Molla, R. (2021). Posting less, posting more, and tired of it all: How the pandemic has changed social media. *Vox Magazine. Recode.* https://www.vox.com/recode/22295131/social-media-use-pandemic-covid-19-instagram-tiktok.

[64] Arendt, D; Blaha, L. (2015). Opinions, influence, and zealotry: a computational study on stubbornness. *Computational and Mathematical Organization Theory.* 20, 184–209. https://link.springer.com/article/10.1007/s10588-015-9181-1.

[65] Meleshevich, K. Schafer, B. Online Information Laundering: The Role of Social Media. *Alliance for Securing Democracy. The German Marshall Fund of the United States.* https://securingdemocracy.gmfus.org/online-information-laundering-the-role-of-social-media/.

[66] Intelligence Community Assessment. (2017). Assessing Russian Activities and Intentions in Recent US Elections. *Officer of the Director of National Intelligence. United States of America.* ICA- 2017-01D. https://www.dni.gov/files/documents/ICA_2017_01.pdf.

[67] Mueller, R. (2019). Report On The Investigation In Russian Interference In The 2016 Presidential Election. U.S. *Department of Justice. Volume 1. CFT 600.8(c)* https://www.justice.gov/archives/sco/file/1373816/download.

multiple cases of this integration have occurred. Leading up to the 2020 U.S. Presidential election, Russian groups tried to plant incriminating documents about then Democratic Nominee Joseph Biden in the preceding months. While many news organizations refused to publish the article due to the inability to corroborate the claims, a few mainstream outlets did, adding legitimacy to the influence campaign.

This legitimization of disinformation can be as simple as reaching a social media influencer who shares it with their followers, elevating its visibility to potentially millions of users. Or it can be nuanced trafficking through multiple legitimate and illegitimate social media platforms, websites, blogs, and algorithms to end up at its target. For example, throughout 2019 and 2020, Russia ran a website called "Peace Data" to influence left leaning citizens in the English-speaking world with the use of known freelance reporters were to add legitimacy to the site [67].[68] These freelance reporters approached on Twitter and asked to write articles on various topics and were paid between $100 and $250 per article. Many of these articles, according to the freelancers, were then spun by the editorial staff to have a particular political angle.

Privacy

While most of the previous techniques may seem like these groups are releasing disinformation on the grand ocean of social media and hoping for something to stick, this couldn't be further from reality. Data brokerage was a $200 billion dollar industry in 2020 and has no signs of slowing down in growth [68].[69] Nearly all social media companies rely on data and advertising, and the sale of both, in order to remain profitable. The adage, "If you're not paying for it, you're not the customer; you're the product" [69][70] has become a staple for the current internet age. As such, a cottage industry has sprouted up around the buying and selling of data, the building of user profiles, demographics of user history, purchases, and website viewings, while predicting their potential actions [70].[71] Hundreds of entities and companies exist solely to buy and sell data for advertising, with many of which are hiding or refusing to show what data they own on U.S. citizens [71].[72]

---

[68] https://www.reuters.com/article/us-usa-election-facebook-russia/duped-by-russia-freelancers-ensnared-in-disinformation-campaign-by-promise-of-easy-money-idUSKBN25T35E.

[69] Knowledge Sourcing Intelligence. (2021).Global Data Broker Market Size, Share, Opportunities, COVID-19 Impact, And Trends By Data Type (Consumer Data, Business Data), By End-User Industry (BFSI, Retail, Automotive, Construction Others), And By Geography—Forecast From 2021. *Report KSI0601611226, Page 131.* https://www.knowledge-sourcing.com/report/global-data-broker-market.

[70] Lewis, A. (2010). User-Driven Content. *Meta Filter Community Blog.* https://www.metafilter.com/95152/Userdriven-discontent#32560467.

[71] Brathwaite, S. (2021). What Does a data broker do? *Security Made Simple.* https://www.securitymadesimple.org/cybersecurity-blog/what-does-a-data-broker-do.

[72] Privacy Rights Clearninghouse. (2021). Data Brokers. https://privacyrights.org/data-brokers.

Social media platforms themselves act more as first part data miners, as they have huge troves of data produced by their users daily. Their proprietary recommendation algorithms claim that they can target specific groups at highly accurate rates [72, 73].[73,74] For example, Facebook allows advertisers to push their content to specific demographics, in specified locations, with selected interests and behaviors, some of which would be illegal in other contexts [74].[75] Facebook has had multiple instances where they allowed people to exclude Black Americans and other minorities from housing advertisements, a violation of the Fair Housing Act [75].[76] Facebook has refused requests to show what advertisements were targeted toward service members, either for fraudulent companies like multi-level marketing schemes or political advertisements. As a result, companies can specifically target U.S. service members or veterans directly, using this highly trusted group to launder misinformation into the mainstream. There is no oversight or regulation of this targeted advertising industry. Individuals and companies are allowed to spend whatever they want with their targeting limited only by the metrics a company. For example, Russia purchased hundreds of thousands of dollars' worth of ads from Facebook leading up to the 2016 U.S. Presidential Election in order to influence its results [76].[77]

Due to limited laws governing data privacy, the onus is placed on the user to protect themselves. This is onerous because, as previously stated, the data being collected can be completely unknown to the end user. In many cases, organizations that do completely other tasks, may be collecting data on a user to sell. In 2020, Avast, a large antivirus company with upwards of 430 million active monthly users, was collecting web browser history across accounts and sessions and selling them on the market [77].[78] Sessions replays act as recreation of a user's action on a Document Object Model, essentially the framework that makes a website or web application, allowing the data collected to be more that strictly clickthrough rates, but to include hovering over certain sections, how they navigate the page itself, and what they seem to focus on [78].[79]

---

[73] Facebook for Business. (2021). Help your ads find the people who will love your business. *Facebook.* https://www.facebook.com/business/ads/ad-targeting.

[74] Twitter For Business. (2021). Twitter Ads Targeting. *Twitter.* https://business.twitter.com/en/advertising/targeting.html.

[75] Facebook for Developers. (2021). Targeting Search, Behaviors. *Facebook.* https://developers.facebook.com/docs/marketing-api/audiences/reference/targeting-search#behaviors.

[76] HUD Public Affairs. (2019). HUD charges Facebook over company's targeted advertising practices with housing discrimination. *U.S. Department of Housing and Urban Development Archives, HUD No. 19–035.* https://archives.hud.gov/news/2019/pr19-035.cfm.

[77] Shane, S. and Goel, V. (2017). Fake Russian Facebook Accounts Bought $100,100 in Political Ads. *The New York Times.* https://www.nytimes.com/2017/09/06/technology/facebook-russian-political-ads.html.

[78] Cox, J. (2020). Leaded Documents Expose the Secretive Market for your Browsing Data. *Vice Magazine.* https://www.vice.com/en/article/qjdkq7/avast-antivirus-sells-user-browsing-data-investigation.

[79] Fullstory. (2021). The Definitive guide to session replay. *Fullstory Learning Center.* https://www.fullstory.com/resources/the-definitive-guide-to-session-replay.

Furthermore, even when users turn off their locational data within a specified app, they are still being tracked through their phone through their Mobile Advertising ID [79].[80] These MAIDs, as they are commonly referred to, allow cellular phone manufacturers to tailor aids to you while not linking personally identifiable information to the person who owns and uses the phone. However, like many other issues with privacy, in practice MAIDs are not anonymous, and there is an industry specifically for unmasking MAIDs and connecting them to real people [80].[81] Being able to track individuals by their device opens several possibilities for near peers to build a pattern of life, directly influence them, or depending on their web content, compromise them.

One of the most notorious cases of this kind of targeted influence is Cambridge Analytica and their scandals involving the 2016 U.S. Presidential Elections and the British EU Referendum. Cambridge Analytica claimed to be able to microtarget users to greater political engagement and voter turnout. The Company, which was a subsidiary of a private intelligence company, used data collected for academic research to microtarget U.S. and U.K. citizens for specific political movements at scale [81].[82] They also incorporated psychographic information, alleging in their marketing materials that they could psychologically target individuals. Whistleblower Christopher Wylie also claimed ties between Cambridge Analytica and Russian government intermediaries in a Senate Judiciary hearing [82].[83] While many users understood that their data might be advertised in products, they may not have been aware that they could potentially be used by foreign countries to cause political turmoil.

[80] Ropek, L. (2021). 'Anonymous' Mobile Advertising IDs Aren't So Anonymoust, And They Are Everywhere. *Gizmondo.* https://gizmodo.com/anonymous-mobile-advertising-ids-aren-t-so-anonymous-1847292216.

[81] Cox, J. (2021). Inside the Industry That Unmasks People at Scale. *Vice Magazine.* https://www.vice.com/en/article/epnmvz/industry-unmasks-at-scale-maid-to-pii.

[82] Kang, C and Frenkel, S. (2018). Facebook Says Cambridge Analytica Harvested Data of Up to 87 Million Users. *The New York Times.* https://www.nytimes.com/2018/04/04/technology/mark-zuckerberg-testify-congress.html.

[83] Wylie, C. (2018) In the Matter of Cambridge Analytica and Other Related Issues. *Written Statement to the United States Senate Committee on the Judiciary.* https://www.judiciary.senate.gov/imo/media/doc/05-16-18%20Wylie%20Testimony.pdf.

Deplatforming and Forced Migration

Following the fallout around Cambridge Analytica and foreign influence campaigns interference in U.S. Politics, a number of social media companies began working through the problem [83–85].[84,85,86] Many used the tactic of removing the content from suspected foreign accounts, or groups infiltrated by foreign entities, which were typically bot or sock puppets. While each platform had their own methodology for identifying accounts or groups, they generally kept that methodology secret to the public. Certain platforms, like Twitter [83],[87] were open about what they removed and archived it for use by other academics and researchers, other platforms kept it a secret and closed ranks. This occurred while simultaneously removing real users who violated that platforms terms of service, typically around inciting violence and hate speech [86, 87],[88,89] or general public safety following the COIVD19 Global Pandemic [88].[90] This removal of real people, commonly referred to as deplatforming, had some short-term success in reducing disinformation and misinformation on the platform [89].[91]

Studies done by the iDrama Lab, led by Jeremy Blackburn of Binghamton University, on two deplatformed subreddits showed that user participation on the platform will decrease after a deplatforming event, reduce the amount of new users, and reduce the amount of overall content [90].[92] A similar study by researchers out of Georgia Institute of Technology on two separate subreddits confirmed the iDrama research, as well as showed that rates of hate speech declined by the users who remained

[84] Twitter. (2021). Information Operations. *Twitter Transparency.* https://transparency.twitter.com/en/reports/information-operations.html.

[85] Facebook. (2021). The State of Influence Operations 2017–2020. *Facebook Threat Report.* https://about.fb.com/wp-content/uploads/2021/05/IO-Threat-Report-May-20-2021.pdf.

[86] Reddit Administrative Staff. (2019). Suspected Campaign from Russia on Reddit. *Reddit Security.* https://www.reddit.com/r/redditsecurity/comments/e74nml/suspected_campaign_from_russia_on_reddit/.

[87] Twitter. (2021). Information Operations. *Twitter Transparency.* https://transparency.twitter.com/en/reports/information-operations.html.

[88] Huffman, S. (2020). Update to Our Content Policy. *Reddit Announcements.* https://www.reddit.com/r/announcements/comments/hi3oht/update_to_our_content_policy/.

[89] Twitter. (2021). Permanent suspension of @realDonaldTrump. *Twitter.* https://blog.twitter.com/en_us/topics/company/2020/suspension.

[90] Facebook. Taking Actions Against Misinformation Across Our Apps. *Facebook.* https://www.facebook.com/combating-misinfo.

[91] Dwoskin, E. and Timberg, C. (2021). Misinformation Dropped Dramatically the week after Twitter banned Trump and some allies. *The Washington Post.* https://www.washingtonpost.com/technology/2021/01/16/misinformation-trump-twitter/.

[92] Ribeiro, MH et all. (2021). Do Platform Migrations Compromise Content Moderations? Evidence from r/The Donald and r/Incels. *24*th *ACM Conference on Computer-Supported Cooperative Work and Social Computing* https://arxiv.org/pdf/2010.10397.pdf.

on the platform [91].[93] For the platforms themselves, this serves as a positive as it removes some of the malevolent forces within them, and shows they have an active moderation and environment free from issues that might lose advertising.

While deplatforming has some benefits, it also has some detriments. One of the biggest is that it can have the tendency to draw attention to a group or movement that hasn't gained mainstream attention [92].[94] While Reddit banned the Qanon Conspiracy for violent rhetoric in 2018, it grew continuously through 2020, reaching its height following the U.S. Presidential Election [93],[95] potentially due in part through the media coverage undermining that deplatforming effort. Adding the taboo of removal from a major social media can push a conspiracy, particularly one centered around persecution, to critical mass [94].[96] While deplatforming will remove potential problems from a specific platform, and potentially make things better within it, it will not remove the ideas and narrative behind the problem.

If users are removed from a platform, or that platform itself is deplatformed from a host, the users within a group will typically attempt to migrate to another more open platform [95].[97] This migration will have multiple results. The first are beneficial, in that typically there isn't a 1 to 1 movement, generally losing active members along the way. The group will also likely not have as large of an audience to recruit from as they would on a large, corporate social media platform [90].[98] Typically, the benefits end there, and instead become benefits for foreign influence campaigns. Users that migrate typically become more insular and radicalized [90].[99] Depending on the political leaning of the group, and the limited moderation, this can be a dangerous combination for political violence. Parler, following mass exodus of conservative users from Twitter, grew substantially through 2020. However, following the Capitol Insurrection on January 6th, 2021, Parler lost its hosting through Amazon Web Services.

---

[93] Chandrasekharan, E et all. (2017). You Can't Stay Here: The Efficacy of Raddit's 2015 Ban Examined Through Hate Speech. *Proceedings of the ACM on Human–Computer Interaction, Vol1, article 31. Pages 1–22* https://dl.acm.org/doi/10.1145/3134666.

[94] Aliapoulious, M et al. (2021). An Early Look at the Parler Online Social Network. *Proceedings of the International AAAI Conference on Web and Social Media, 15 pages 943–951.* https://arxiv.org/pdf/2101.03820.pdf.

[95] Dickson, EJ. (2021). The Qanon Community Is in Crisis—But On Telegram, It's Also Growing. *Rolling Stone Magazine.* https://www.rollingstone.com/culture/culture-news/qanon-telegram-channels-increase-1117869/.

[96] Nickas, J; Isaac, M; Frenkel, S. (2021). Millions Flock to Telegram and Signal as Fear Grows Over Big Tech. *The New York Times.* https://www.nytimes.com/2021/01/13/technology/telegram-signal-apps-big-tech.html.

[97] Johnson NF et all. (2019). Hidden resilience and adaptive dynamics of the global online hate ecology. *Nature 573, pages 261–265.* https://www.nature.com/articles/s41586-019-1494-7.

[98] Ribeiro, MH et all. (2021). Do Platform Migrations Compromise Content Moderations? Evidence from r/The Donald and r/Incels. *24*th *ACM Conference on Computer-Supported Cooperative Work and Social Computing* https://arxiv.org/pdf/2010.10397.pdf.

[99] Ribeiro, MH et all. (2021). Do Platform Migrations Compromise Content Moderations? Evidence from r/The Donald and r/Incels. *24*th *ACM Conference on Computer-Supported Cooperative Work and Social Computing* https://arxiv.org/pdf/2010.10397.pdf.

This led Parler to turn to DDoS-Guard for these services, a company with ties to the Russian Federation, who's owners are based in Rostov-on-Don, Russia [96].[100]

These smaller offshoots, also have more questionable security practices compared to their larger corporate counterparts. Activists attempted to catalog Parler posts following the Capitol Insurrection and found a vulnerability that allowed them to pull all data from the platform at an admin level, including deleted posts, messages, and locational data [97].[101] Similarly, Gab, a platform who hosted the 2018 Tree of Life synagogue shooter [98],[102] was hacked the next month, leaking private messages and content similar to that of the Parler leak [99].[103] Both of these platforms include U.S. congress members and U.S. Senators, along with a slew of state legislators, mayors, and governors. Another study done by the iDrama lab, in cooperation with Boston University's SecLAB, showed that many users deplatformed go to these alternative platforms only to become more vitriolic, although losing much of the size of their following in doing so [100].[104]

Generative Media

New machine learning algorithms that produce realistic images, videos, audio, and text are being created every day. These algorithms are generally grouped under the name of Generative Adversarial Networks [101],[105] for the way that two algorithms work against each other, like a counterfeiter against law enforcement, to generate realistic media. The most notable example is the titular deepfakes, named as a portmanteau of deep neural networks that create fake media. These videos, as they were found, typically replaced women in adult films with the faces of celebrities [102].[106] Over the course of 2019 and 2020, these algorithms raised great concern

[100] Menn, J; Li, K; Culliford, E. (2021). Parler partially reappears with support from Russian technology firm. *Reuters Internet News.* https://www.reuters.com/article/us-usa-trump-parler-russia/parler-partially-reappears-with-support-from-russian-technology-firm-idINKBN29N23N.

[101] Holt, J; Brookie, G, Brooking, E. (2021). Fast Thinking: How the Capitol riot was coordinated online. *The Atlantic Council.* https://www.atlanticcouncil.org/content-series/fastthinking/fast-thinking-how-the-capitol-riot-was-coordinated-online/.

[102] Saldiva, G; Van Sant, S; Bowman, E. (2018). Suspect Charged With 29 Federal Counts in Pittsburgh Synagogue Massacre. *National Public Radio.* https://www.npr.org/2018/10/27/661347236/multiple-casualties-in-shooting-near-pittsburgh-synagogue.

[103] Greenburg, A. (2021). Far-Right Platform Gab Has Been Hacked—Including Private Data. *Wired Magazine.* https://www.wired.com/story/gab-hack-data-breach-ddosecrets/.

[104] Ali, S et all. (2021). Understanding the Effect of Deplatforming on Social Networks. WebSci'12: 13th ACM Web Science Conference pages 187–195. https://seclab.bu.edu/people/gianluca/papers/deplatforming-websci2021.pdf.

[105] Goodfellow, I et all. (2014). Generative Adversarial Nets. Proceedings of the 27th *International Conference on Neural Information Processing Systems, Volume 2, pages 2672–2680.* https://arxiv.org/pdf/1406.2661.pdf.

[106] Cole, S. (2017). AI-Assisted Fake Porn Is Here. *Vice Magazine.* https://www.vice.com/en/article/gydydm/gal-gadot-fake-ai-porn.

among the legal circles [103],[107] government officials [104],[108] defense analysts [105],[109] academia [106],[110] and industry [107][111] for the potential for their use in disinformation campaigns.

Many of these algorithms are being open sourced and shared on programming repositories like github and bitbucket [108, 109],[112,113] open for use by the general public, and foreign adversaries. In fact, many have already been found in the wild. Foreign Intelligence Services have already used deep fakes to add credibility to sock puppet accounts meant for collecting information of U.S. service members and those of the intelligence community on LinkedIn [110].[114] Domestically, U.S. political groups have used sock puppet accounts to weave a narrative leading up to elections on Facebook [111].[115] These technologies have been so democratized that a Pennsylvanian mother sent deepfaked pictures of her daughter's cheerleading rival [112].[116]

---

[107] Citron, D. (2019). The National Security Challenge of Artifical Intelligence, Manipulated Media, and "Deep Fakes." *Prepared Written Testimony and Statement for the Record. House Permanent Select Committee on Intelligence.* https://intelligence.house.gov/uploadedfiles/citron_testimony_for_house_committee_on_deep_fakes.pdf.

[108] Rubio, M and Warner, M. (2019). Rubio, Warner Express Concern Over Growing Threat Posed by Deepfakes. *Joint Statement from Senators Marco Rubio and Mark Warner.* https://www.rubio.senate.gov/public/index.cfm/2019/10/rubio-warner-express-concern-over-growing-threat-posed-by-deepfakes.

[109] Watts, C. (2019). The National Security Challenge of Artifical Intelligence, Manipulated Media, and "Deep Fakes." *Prepared Written Testimony and Statement for the Record. House Permanent Select Committee on Intelligence.* https://intelligence.house.gov/uploadedfiles/clint_watts_-_house_select_committee_on_intelligence_-_ai__deep_fakes_-_13_june_2019.pdf.

[110] Doermann, D. (2019). The National Security Challenge of Artifical Intelligence, Manipulated Media, and "Deep Fakes." *Prepared Written Testimony and Statement for the Record. House Permanent Select Committee on Intelligence.* https://intelligence.house.gov/uploadedfiles/doermann-statement-final.pdf.

[111] Clark, J. (2019). The National Security Challenge of Artifical Intelligence, Manipulated Media, and "Deep Fakes." *Prepared Written Testimony and Statement for the Record. House Permanent Select Committee on Intelligence.* https://intelligence.house.gov/uploadedfiles/clark_deepfakes_sfr.pdf.

[112] Deepfakes. (2019). FaceSwap Manifesto. *Github pages.* https://github.com/deepfakes/faceswap#manifesto.

[113] Radek. (2018). Myfakeapp. *Bitbucket.* https://bitbucket.org/radeksissues/myfakeapp/src/master/.

[114] Boyd. M. (2021). Deepfakes and LinkedIn: malign interference campaigns. *Malwarebytes Labs, Social Engineering.* https://blog.malwarebytes.com/social-engineering/2019/11/deepfakes-and-linkedin-malign-interference-campaigns/.

[115] Collins, B. (2019). Facebook says a pro-Trump media outlet used artificial intelligence to create fake people and push conspiracies. *National Broadcasting Company News.* https://www.nbcnews.com/tech/tech-news/facebook-says-pro-trump-media-outlet-used-artificial-intelligence-create-n1105951.

[116] Associated Press. (2021). Cheerleader's mom accused of making 'deepfakes' of rivals. The Associated Press https://apnews.com/article/pennsylvania-doylestown-cheerleading-0953a60ab3e3452b87753e81e0e77d7f.

These algorithms have a deep amount of potential for abuse by Russia, China, and Iran in disinformation campaigns. As stated above, they are already being used in profile pictures on bot and sock puppet account profiles to add a layer of believability. nVidia Corporation's StyleGAN [113, 114][117,118], and StyleGAN2 [115, 116][119,120] all create detailed and realistic still images that are completely synthetic using a technique called style transfer, which takes characteristics from one image and applies them to another. Both trained models have been released to the public, and both have been used in created accounts.

The reach of GANs and Deepfakes goes much further than that. An algorithm created by Samsung Moscow can create realistic moving images and videos based off a single image of a person [117].[121] With more images, from different angles, the quality will increase substantially, making seamless videos over the top of a staged video. This is a substantial decrease from something like the 'face swap' app that gave way to the initial rounds of deepfakes, which required hundreds if not thousands of images of the target to create a realistic fake. While many news agencies argue the main threat is political, in practice most targets thus far have been people without the means to pushback against [118].[122] This means the targeting of government officials, particularly those vulnerable in overseas assignments [119],[123] for blackmail and extortion [120],[124] which are already commonplace in a foreign intelligence agency's arsenal [121].[125]

---

[117] Kerras, T; Laine, S; Aila T. (2019). A Style-Based Generator Architecture for Generative Adversarial Networks. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).* https://arxiv.org/abs/1812.04948.

[118] nVidia Labs. (2019). Style-Gan Repository. *Github.* https://github.com/NVlabs/stylegan.

[119] Kerras, T et all. (2020). Analyzing and Improving the Image Quality of StyleGAN. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).* https://arxiv.org/abs/1912.04958.

[120] nVidia Labs. (2020). Style-Gan2 Repository. *Github.* https://github.com/NVlabs/stylegan2.

[121] Zakhorav, E et all. (2019) Few-Shot Adversarial Learning of Realistic Neural Talking Head Models. *2019 IEEE/CVF International Conference on Computer Vision (ICCV).* https://arxiv.org/pdf/1905.08233.pdf.

[122] Gosse, C. and Burkell, J. (2020). Politics and porn: how news media characterizes problems presented by deepfakes. *Critical Studies in Media Communications, Volume 30 Issue 5.* https://nca.tandfonline.com/doi/full/10.1080/15295036.2020.1832697?journalCode=rcsm20#.YPrydo5KhPa.

[123] Littell, J. (2019). Don't Believe Your Eyes (or Ears): The Weaponizations of Artificial Intelligence, Machine Learning, And Deepfakes. *War on the Rocks.* https://warontherocks.com/2019/10/dont-believe-your-eyes-or-ears-the-weaponization-of-artificial-intelligence-machine-learning-and-deepfakes/.

[124] Weiss, M. (2019). The Hero Who Betrayed His Country. *The Atlantic.* https://www.theatlantic.com/international/archive/2019/06/estonia-russia-deniss-metsavas-spy/592417/.

[125] Nemtsova, A. (2019). What to Do When the Russian Government Wants to Blackmail You. *The Atlantic.* https://www.theatlantic.com/international/archive/2019/03/russia-government-blackmail-kompromat/585850/.

Couple the visualization with synthetic audio [122],[126] and you have very realistic fakes [123][127] that can be made cheaply, quickly, and distributed through the aforementioned information laundering methodology. Foreign adversaries could cause protests similar to the 2014 Benghazi consulate attack based on fake video or audio. This could be equally worrisome for both domestic and foreign operations, where soft targets could be overwhelmed by unknowing populations upset at fake media. Russia has already expertly orchestrated real world protests using Facebook groups and advertising [124],[128] adding in perfectly crafted fake media could bring that capability to scale.

And while not Generative Adversarial Networks or deepfakes per se, generative text [125][129] stands to expand and automate many processes already in use. OpenAI's Generative Pretrained Transformers (GPT) [126][130] alone has multiple applications already in the wild. Users can create realistic social media bots that pull topics from currently tending topics and create appropriate comments within them. This allows for robust botnets and sock puppet accounts that will circumvent some of the trusted methods for hunting and removing their accounts. Other media can be created from these models as well. Meme generators based of a string of text can create visual media [127][131] that can be more pervasive and influential than simple text [128].[132] Even mainstream news agencies are utilizing model generated text to fill out, or in other cases entirely write articles on a specific subject with minimal inputs [129].[133] Given Russians preponderance to create entire news sites and paying real journalists for legitimacy, they could potentially automate much of the process for a similar effect.

---

[126] Tung, L. (2016). Adobe's VoCo voice project: Now you really can put words in someone else's mouth. *ZDnet.* https://www.zdnet.com/article/adobes-voco-voice-project-now-you-really-can-put-words-in-someone-elses-mouth/.

[127] Dessa Corporation. (2019). RealTalk: We Recreated Joe Rogan's Voice Using Artificial Intelligence. YouTube. https://youtu.be/DWK_iYBl8cA.

[128] Kosoff, M. (2017). How Russia Secretly Orchestrated Dozens of U.S. Protests. *Vanity Fair.* https://www.vanityfair.com/news/2017/10/how-russia-secretly-orchestrated-dozens-of-us-protests.

[129] Kawthekar, P; Rewari, R; Bhooshan, S. (2017). Evaluating Generative Models for Text Generation. *Stanford University.* https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1174/reports/2737434.pdf.

[130] Brown, T et all. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems 33.* https://arxiv.org/abs/2005.14165.

[131] Anand, A. (2020). Another 10 gems of GPT-3. *Dev.to Blog.* https://dev.to/amananandrai/another-10-gems-of-gpt-3-2639.

[132] Huntington, H. (2017). The affects and Effect of Internet Memes: Assessing Perceptions and Influence of Online User-Generated Poltical Discourse as Media. *Colorado State University Department of Journalism and Media Communication.* https://mountainscholar.org/bitstream/handle/10217/183936/Huntington_colostate_0053A_14303.pdf.

[133] Hutson, M. Robo-writers: the rise and risks of language generating AI. *Nature News Feature.* https://www.nature.com/articles/d41586-021-00530-0.

Data Poisoning and Contamination

As automation of disinformation takes hold, processes to find and remove it will also turn to automation to counter it. As such, much like cyber security, an arms race will begin between those nations who look to use online spaces for the spread of disinformation, and those who look to remove it. In order to counter the sheer volume of disinformation being created, and forensically determine authenticity of various pieces of media, algorithmic methods will need to be used, as human analysts alone will not be able to process the volume, variety, and veracity of the data. This, in turn, will lead to innovative ways to get around these algorithmic filters.

The most likely course of action is to develop adversarial machine learning to disrupt and outright disable the algorithms trying to identify various methods, media, and pathways of disinformation. Typically, this is referred to as data poisoning [130][134] or, if the model is already created and running, data contamination [131].[135] Data poisoning serves as an adversarial machine learning attack that focuses on adding data that the algorithm learns from at the time of model creation. It attempts to bias the decisions made by the algorithmic model so that selected data goes through undetected. For a clearer example, if Russia wanted to have more believable accounts, they could use deepfake images from a GAN to make profile pictures. If the platform, or another nation wanted to find these accounts, they may train a computer vision algorithm to look for them [132].[136] Given the game of cat and mouse, the Russian creating the media may have anticipated some form of intervention, and instead overlaid their images with a small marker, only visible to the computer, but not to the human who labels the data. Thus, the algorithm made to identify and mark these suspicious accounts will most likely be accounts that are identified by the placed marker, leaving other accounts, that were not meant to be found easily, to pass through the algorithmic net.

Comparatively speaking, adversarial data contamination works similarly, but instead of the algorithm being built off of the data, it's being updated by new data after having been already trained. Overall, the methodology would remain the same, but there would be a larger number of accounts that need to be identified in order to slowly shift the bias in the direction our adversaries would like. This type of attack is much more likely, as platforms and others have already begun to algorithmically filter out fake accounts and synthetic media.

---

[134] Xiao, H et all. (2015). Is Feature Selection Secure against Training Data Poisoning. *Proceedings in the 3nd International Conference on Machine Learning.* http://proceedings.mlr.press/v37/xiao15.pdf.

[135] Xiao, H. et all (2015). Support vector machines under adversarial label contamination. *Neurocomputing Volume 160, pages 53–62.* https://www.sec.in.tum.de/i20/publications/support-vector-machines-under-adversarial-label-contamination/file/main-revision.pdf.

[136] Nguyen, T et all. (2019). Deep Learning for Deepfakes Creation and Detection: A Survey. *Cornell University Arxiv.* https://arxiv.org/pdf/1909.11573.pdf.

Conclusion

More and more countries are heavily leaning into online disinformation to further their political goals both domestically and abroad. Many of which have put information at the forefront of their military and diplomatic campaigns to gain the advantage in this ever more connected world. Data is readily available for potentially building psychological profiles of groups at scale, or individually for more precise influencing. This leads to many different dangers from government and military service members to the very democratic institutions they uphold. Facts can be drowned out from the conversation to allow fringe beliefs to seem as though they are commonplace.

Russia and China have decades of online manipulation in their repertoire already being turned onto the United States, and the greater ideals of western democracies. Daron Acemoglu and James A. Robinson, authors of "The Narrow Corridor," describe the fragile balance between state and society that allow liberal democracy to avoid falling into the alternatives of lawlessness or authoritarianism [133].[137] If the state becomes too strong, it risks despots removing liberties from the people, if the society becomes too strong, it falls into a state of anarchy, where the state's ability to protect the populace from groups within fails. Long term influence campaigns can tilt the balance of power either way, allowing for the house of cards, that is our freedom, to come tumbling down.

To keep democracy from falling to tatters, we must institute better protection of our privacy, in particular our data, while our national defense pushes back against outside actors who look to break it down. Without full-fledged, proactive countering of our near peers, we will be continuously playing catchup, fighting off our back foot merely trying to keep our heads above water. Without change to our current stature, we will inevitably see greater fissures in our society, and less trust in our state. Without the ability of citizens to actively participate in their government, through free and fair elections, and build consensus across different belief structures and groups of citizenries, democracy cannot function in a state that is accepted. It will give way to other, harsher forms of government, that will see these failures as an opportunity to seize power at the expense of the rights of many of its citizens.

References

1. Mattis, P. (2018). China's Three Warfares in Perspective. *War on the Rocks. Special Series: Ministry of Truth.* https://warontherocks.com/2018/01/chinas-three-warfares-perspective/.
2. Cunningham, C. (2020). A Russian Federation Information Warfare Primer. *Then Henry M. Jackson School of International Studies. Washington University.* https://jsis.washington.edu/news/a-russian-federation-information-warfare-primer/.

[137] Acemoglu, D; Robinson, J. (2019). The Narrow Corridor. States, Societies, and the Fate of Liberty. *Penguin Random House Publishing.*

3. Strick, B. (2020). COVID-19 Disinformation: Attempted Influence in Disguise. *Australian Strategic Policy Institute. International Cyber Policy Center.* https://www.aspi.org.au/report/covid-19-disinformation.
4. Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. (2017). Generative Adversarial Networks: An Overview. IEE-SPM. April 2017. https://arxiv.org/pdf/1710.07035.pdf.
5. Whittaker, L., Letheren, K., & Mulcahy, R. (2021). The Rise of Deepfakes: A conceptual framework and research agenda for marketing. https://doi.org/10.1177/1839334921999479.
6. https://arxiv.org/abs/1810.04805.
7. Bradshaw, S., Bailey, H., & Howard, P. (2020). 2020 Global Inventory of Organized Social Media Manipulation. *Computational Propaganda Research Project. Oxford Internet Institute. Oxford University.* https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/127/2021/02/CyberTroop-Report20-Draft9.pdf.
8. https://diamond-democracy.stanford.edu/speaking/lectures/what-democracy.
9. Bastos, M., & Mercea, D. (2018). The public accountability of social platforms: lessons from a study on bots and trolls in the Brexit campaign. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, *376*(2128), 20180003. https://doi.org/10.1098/rsta.2018.0003.
10. National Intelligence Council. (2021). Foreign threats to the 2020 US federal elections. *Intelligence Community Assessment.* https://www.dni.gov/files/ODNI/documents/assessments/ICA-declass-16MAR21.pdf.
11. Wood, D., McMinn, S., & Feng, E. (2019). China Used Twitter to Disrupt Hong Kong Protests, but Efforts Began Years Earlier. *National Public Radio.* https://www.npr.org/2019/09/17/758146019/china-used-twitter-to-disrupt-hong-kong-protests-but-efforts-began-years-earlier.
12. Uyghur Human Rights Project. (2020). The Happiest Muslims in the World: Disinformation, Propaganda, and the Uyghur Crisis. *Uyghur Human Rights Project*. https://docs.uhrp.org/pdf/Disinformation_Propganda_and_the_Uyghur_Crisis.pdf.
13. Bernard, R., Bowsher, G., Sullivan, R., & Gibson-Fall, F. (2021). Disinformation and Epidemics: Anticipating the Next Phase of Biowarfare. *Health Security. Volume 19, Number 1*. https://doi.org/10.1089/hs.2020.0038.
14. Giles, C., & Bhat, U. (2020). Nagoro-Karabakh: The Armenian-Azeri 'Information Wars'. *BBC Reality Checks and Anti-disinformation Unit.* https://www.bbc.com/news/world-europe-54614392.
15. Beech, H. (2021). 'Now We Are United': Myanmar's Ethnic Divisions Soften after Coup. *New York Times.* https://www.nytimes.com/2021/04/30/world/asia/myanmar-ethnic-minority-coup.html.
16. McKinsey and Company. (2021). Covid-19: Implications for Business. *McKinsey and Company. Executive Briefing.* https://www.mckinsey.com/business-functions/risk-and-resilience/our-insights/covid-19-implications-for-business.
17. Center for System Sciences and Engineering (CSSE). (2021). COVID-19 Dashboard. *Johns Hopkins University. ARCGIS.* https://www.arcgis.com/apps/dashboards/bda7594740fd40299423467b48e9ecf6.
18. Blomberg, B., Mohn, K., & Brokstad, K. A. et al. (2021). Long COIVD in a prospective cohort of home-isolated patients. *Nature Medicine.* https://www.nature.com/articles/s41591-021-01433-3.
19. Centers for Disease Control and Prevention. (2020). CDC Calls on Americans to wear mask to prevent COIVD-19 Spread. *Centers for Disease Control and Prevention.* https://www.cdc.gov/media/releases/2020/p0714-americans-to-wear-masks.html.
20. Loucaides, D., & Perrone, A. (2021). Germany's COVID sceptics fueled by Russian media and far-Right Conspiracies. *openDemocracy.* https://www.opendemocracy.net/en/germanys-covid-sceptics-fuelled-by-russian-media-and-far-right-conspiracies/.
21. Benson, T. (2020). Trolls and bots are flooding social media with disinformation encouraging states to end quarantine. *Business Insider. Tech.* https://www.businessinsider.com/trolls-bots-flooding-social-media-with-anti-quarantine-disinformation-2020-4.

22. Emmot, R. (2020). Russia deploying coronavirus disinformation to sow panic in West, EU document says. *Reuters.* https://www.reuters.com/article/us-health-coronavirus-disinformation/russia-deploying-coronavirus-disinformation-to-sow-panic-in-west-eu-document-says-idUSKBN21518F.
23. Miller, H. (2021). Russia's Anti-Vaccine Propaganda is Tatamount to a Declaration of War. *Center for Medical Economic and Innovation.* https://medecon.org/russias-anti-vaccine-propaganda-is-tantamount-to-a-declaration-of-war/.
24. Chen, E. (2021). Chinese COVID-19 Misinformation A Year Later. *The Jamestown Foundation Global Research and Analysis.* https://jamestown.org/program/chinese-covid-19-misinformation-a-year-later/.
25. Reuters Staff. (2020). China sends medical supplies, experts to help Italy battle coronavirus. *Reuters Healthcare and Pharma.* https://www.reuters.com/article/us-health-coronavirus-italy-respirators/china-sends-medical-supplies-experts-to-help-italy-battle-coronavirus-idUSKBN2101IM.
26. Mallapaty, S. (2021). China's COVID vaccines are going global—but questions remain. *Nature Magazine* https://www.nature.com/articles/d41586-021-01146-0.
27. Kinetz, E. (2021). Anatomy of a conspiracy: With COVID, China took leading role. *Associated Press.* https://apnews.com/article/pandemics-beijing-only-on-ap-epidemics-media-122b73e134b780919cc1808f3f6f16e8.
28. Taylor, P. (2003). Munitions of the mind: A history of propaganda. *Manchester University Press.* https://www.jstor.org/stable/j.ctt155jd69.
29. Edwards, M. (1994). Printing, Propaganda, and Martin Luther. *Berkeley: University of California Press.* https://publishing.cdlib.org/ucpressebooks/view?docId=ft3q2nb278;chunk.id=0;doc.view=print.
30. https://publishing.cdlib.org/ucpressebooks/view?docId=ft3q2nb278;chunk.id=0;doc.view=print.
31. https://www.ushmm.org/collections/bibliography/nazi-propaganda-1.
32. Jowett, G., & O'Donnell, V. (2012). *Propaganda and Persuasion*, 5th Edn. Sage Publications.
33. https://watch.eventive.org/m4df2021/play/6074567c86f143003e0cbd99/6074697b69154f003086bb13.
34. Broderick, R. (2014). Activists Are Outing Hundreds Of Twitter Users Believed To Be 4chan Trolls Posing As Feminist. *BuzzFeed News.* https://www.buzzfeednews.com/article/ryanhatesthis/your-slip-is-showing-4chan-trolls-operation-lollipop.
35. Ferrara E. (2020). Bots, Elections, and Social Media: A Brief Overview. In K. Shu, S. Wang, D. Lee, & H. Liu (eds) *Disinformation, misinformation, and fake news in Social Media.* Lecture Notes in Social Networks. Springer, Cham. https://doi.org/10.1007/978-3-030-42699-6_6.
36. Bakir, V., & McStay, A. (2018). Fake news and the economy of emotions. *Digital Journalism, 6*(2), 154–175. https://doi.org/10.1080/21670811.2017.1345645
37. Anderson, J., Rainie, L., Vogels, e. (2021). Expers Say the 'New Normal' in 2025 Will Be Far More Tech-Driven, Presenting More Big Challenges. Pew Research Center. *Emerging Technologies. Future of the Internet.* https://www.pewresearch.org/internet/2021/02/18/experts-say-the-new-normal-in-2025-will-be-far-more-tech-driven-presenting-more-big-challenges/.
38. Polysanskaya, A., Krivov, A., &Lomko, I. (2003). The Virtual Eye of Big Brother. *Vestnik Online.* http://www.vestnik.com/issues/2003/0430/win/polyanskaya_krivov_lomko.htm.
39. Elder, M. (2012). Polishing Putin: hacked emails suggest dirty tricks by Russian youth group. *The Guardian.* https://www.theguardian.com/world/2012/feb/07/putin-hacked-emails-russian-nashi.
40. Bandurski, D. (2008). China's Guerrilla War for the Web. *China Media Project. China Internet Research Conference.* https://chinamediaproject.org/2008/07/07/feer-chinas-guerrilla-war-for-the-web/.
41. King, G., Pan, J., & Roberts, M. (2017). How the Chinese Government Fabricates Social Media Post for Strategic Distraction, not Engaged Argument. *American Political Science Review, vol. 111, pg. 484–501.* https://gking.harvard.edu/files/gking/files/50c.pdf.

42. Lock, I., & Ludolph, R. (2019). Organizational propaganda on the Internet: A systematic review. *Public Relations Inquiry. vol 9, issue 1, pg 103–127.* https://doi.org/10.1177/2046147X19870844.
43. Kenny, C., & Bergman, M. (2019). Understanding and Combating Russian and Chinese Influence Operations. *Center for American Progress, Foreign Policy and Security.* https://www.americanprogress.org/issues/security/reports/2019/02/28/466669/understanding-combating-russian-chinese-influence-operations/.
44. Select Committee on Intelligence. United States Senate. (2018). Russian Active Measures Campaigns and Interference in the 20216 U.S. Election Volume 2: Russia's Use of Social Media with Additional Views. *116*th *Congress. Senate. 1*st *Session. Report 116-XX.* https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume2.pdf.
45. Meserole, C. (2018). How misinformation spreads on social media—and what to do about it. *Brookings Institute.* https://www.brookings.edu/blog/order-from-chaos/2018/05/09/how-misinformation-spreads-on-social-media-and-what-to-do-about-it/.
46. Arjomand, N. (2019). Information Laundering and Globalized Media—Part I: The Problem. *Center for International Media Assistance.* https://www.cima.ned.org/blog/information-laundering-and-globalized-media-part-i-the-problem/.
47. Sherman, J. (2021). Federal Privacy Rules Must Get "Data Broker" Definitions Right. Lawfare. *Consumer Data.* https://www.lawfareblog.com/federal-privacy-rules-must-get-data-broker-definitions-right.
48. The National Law Review. (2019). Data Scraping Survives! (At Least for Now) Key Takeaways from 8th Circuit Ruling on the HIQ vs. LinkedIn Case. *The National Law Review.* https://www.natlawreview.com/article/data-scraping-survives-least-now-key-takeaways-9th-circuit-ruling-hiq-vs-linkedin.
49. Diamond, L. (2004). What is Democracy. *Lecture to Hilla University for Humanistic Studies.* https://diamond-democracy.stanford.edu/speaking/lectures/what-democracy.
50. Google Developers. (2021). Introduction to Recommendation Systems. *Google.* https://developers.google.com/machine-learning/recommendation.
51. Hardesty, L. (2019). The history of Amazon's recommendation algorithm. *Amazon Science.* https://www.amazon.science/the-history-of-amazons-recommendation-algorithm.
52. Ilic, A. & Kabiljo, M. (2015) Recommending items to more than a billion people. *Facebook Engineering, Core Data, ML Applications.* https://engineering.fb.com/2015/06/02/core-data/recommending-items-to-more-than-a-billion-people/.
53. Oracle. (2020). Introduction to Recommendation Engines. *Oracle Cloud Infrastructure Data Science.* https://www.oracle.com/a/ocom/docs/introduction-to-recommendation-engines-wp.pdf.
54. Chaney, A., Stewart, B., & Engelhardt, B. (2018). How Algorithmic Confounding in Recommendation Systems Increases Homogeneity and Decreases Utility. *Twelfth ACM Conference on Recommender Systems (RecSys '18).* https://scholar.princeton.edu/bstewart/publications/how-algorithmic-confounding-recommendation-systems-increases-homogeneity-and.
55. Fisher, M., & Taub, A. (2019). How YouTube Radicalized Brazil. *New York Times.* https://www.nytimes.com/2019/08/11/world/americas/youtube-brazil.html.
56. Solsman, J. (2018). YouTube's AI is the puppet master over most of what you watch. *CNET*. https://www.cnet.com/news/youtube-ces-2018-neal-mohan/.
57. Girginova, K. (2017). Hojacking Heads and Hashtags. *Global Journal, vol 10, Issue 56.* https://globalejournal.org/global-e/august-2017/hijacking-heads-hashtags.
58. Huffman, S. (2016). TIFU by editing some comments and creating an unnecessary controversy. *Reddit Announcements.* https://www.reddit.com/r/announcements/comments/5frg1n/tifu_by_editing_some_comments_and_creating_an/.
59. Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science, 359*(6380), 1146–1151. https://science.sciencemag.org/content/359/6380/1146.
60. Meleshevich, K., & Schafer, B., Online Information Laundering: The Role of Social Media. *Alliance for Securing Democracy. The German Marshall Fund of the United States.* https://securingdemocracy.gmfus.org/online-information-laundering-the-role-of-social-media/.

61. Center for Countering Digital Hate. (2021). This Disinformation Dozen. *Center for Countering Digital Hate.* https://www.counterhate.com/disinformationdozen.
62. Parasocial interaction. https://doi.org/10.1093/oi/authority.20110803100305809.
63. Molla, R. (2021). Posting less, posting more, and tired of it all: How the pandemic has changed social media. *Vox Magazine. Recode.* https://www.vox.com/recode/22295131/social-media-use-pandemic-covid-19-instagram-tiktok.
64. Arendt, D., & Blaha, L. (2015). Opinions, influence, and zealotry: a computational study on stubbornness. *Computational and Mathematical Organization Theory, 20*, 184–209. https://doi.org/10.1007/s10588-015-9181-1.
65. Intelligence Community Assessment. (2017). Assessing Russian Activities and Intentions in Recent US Elections. *Officer of the Director of National Intelligence. United States of America.* ICA- 2017–01D. https://www.dni.gov/files/documents/ICA_2017_01.pdf.
66. Mueller, R. (2019). Report On The Investigation In Russian Interference In The 2016 Presidential Election. U.S. *Department of Justice. Volume 1. CFT 600.8(c)* https://www.justice.gov/archives/sco/file/1373816/download.
67. https://www.reuters.com/article/us-usa-election-facebook-russia/duped-by-russia-freelancers-ensnared-in-disinformation-campaign-by-promise-of-easy-money-idUSKBN25T35E.
68. Knowledge Sourcing Intelligence. (2021).Global Data Broker Market Size, Share, Opportunities, COVID-19 Impact, And Trends By Data Type (Consumer Data, Business Data), By End-User Industry (BFSI, Retail, Automotive, Construction Others), And By Geography—Forecast From 2021. *Report KSI0601611226, Page 131.* https://www.knowledge-sourcing.com/report/global-data-broker-market.
69. Lewis, A. (2010). User-Driven Content. *Meta Filter Community Blog.* https://www.metafilter.com/95152/Userdriven-discontent#32560467.
70. Brathwaite, S. (2021). What Does a data broker do? *Security Made Simple.* https://www.securitymadesimple.org/cybersecurity-blog/what-does-a-data-broker-do.
71. Privacy Rights Clearninghouse. (2021). Data Brokers. https://privacyrights.org/data-brokers.
72. Facebook for Business. (2021). Help your ads find the people who will love your business. *Facebook.* https://www.facebook.com/business/ads/ad-targeting.
73. Twitter For Business. (2021). Twitter Ads Targeting. *Twitter.* https://business.twitter.com/en/advertising/targeting.html.
74. Facebook for Developers. (2021). Targeting Search, Behaviors. *Facebook.* https://developers.facebook.com/docs/marketing-api/audiences/reference/targeting-search#behaviors.
75. HUD Public Affairs. (2019). HUD charges Facebook over company's targeted advertising practices with housing discrimination. *U.S. Department of Housing and Urban Development Archives, HUD No. 19–035.* https://archives.hud.gov/news/2019/pr19-035.cfm.
76. Shane, S., & Goel, V. (2017). Fake Russian Facebook Accounts Bought $100,100 in Political Ads. *The New York Times.* https://www.nytimes.com/2017/09/06/technology/facebook-russian-political-ads.html.
77. Cox, J. (2020). Leaded documents expose the secretive market for your browsing data. *Vice Magazine.* https://www.vice.com/en/article/qjdkq7/avast-antivirus-sells-user-browsing-data-investigation.
78. Fullstory. (2021). The Definitive guide to session replay. *Fullstory Learning Center.* https://www.fullstory.com/resources/the-definitive-guide-to-session-replay.
79. Ropek, L. (2021). 'Anonymous' Mobile Advertising IDs Aren't So Anonymoust, And They Are Everywhere. *Gizmondo.* https://gizmodo.com/anonymous-mobile-advertising-ids-aren-t-so-anonymous-1847292216.
80. Cox, J. (2021). Inside the Industry That Unmasks People at Scale. *Vice Magazine.* https://www.vice.com/en/article/epnmvz/industry-unmasks-at-scale-maid-to-pii.
81. Kang, C., & Frenkel, S. (2018). Facebook Says Cambridge Analytica Harvested Data of Up to 87 Million Users. *The New York Times.* https://www.nytimes.com/2018/04/04/technology/mark-zuckerberg-testify-congress.html.
82. Wylie, C. (2018) In the Matter of Cambridge Analytica and Other Related Issues. *Written Statement to the United States Senate Committee on the Judiciary.* https://www.judiciary.senate.gov/imo/media/doc/05-16-18%20Wylie%20Testimony.pdf.

83. Twitter. (2021). Information Operations. *Twitter Transparency.* https://transparency.twitter.com/en/reports/information-operations.html.
84. Facebook. (2021). The State of Influence Operations 2017–2020. *Facebook Threat Report.* https://about.fb.com/wp-content/uploads/2021/05/IO-Threat-Report-May-20-2021.pdf.
85. Reddit Administrative Staff. (2019). Suspected Campaign from Russia on Reddit. *Reddit Security.* https://www.reddit.com/r/redditsecurity/comments/e74nml/suspected_campaign_from_russia_on_reddit/.
86. Huffman, S. (2020). Update to Our Content Policy. *Reddit Announcements.* https://www.reddit.com/r/announcements/comments/hi3oht/update_to_our_content_policy/.
87. Twitter. (2021). Permanent suspension of @realDonaldTrump. *Twitter.* https://blog.twitter.com/en_us/topics/company/2020/suspension.
88. Facebook. Taking Actions Against Misinformation Across Our Apps. *Facebook.* https://www.facebook.com/combating-misinfo.
89. Dwoskin, E., & Timberg, C. (2021). Misinformation Dropped Dramatically the week after Twitter banned Trump and some allies. *The Washington Post.* https://www.washingtonpost.com/technology/2021/01/16/misinformation-trump-twitter/.
90. Ribeiro, M. H. et al. (2021). Do Platform Migrations Compromise Content Moderations? Evidence from r/The Donald and r/Incels. *24*th *ACM Conference on Computer-Supported Cooperative Work and Social Computing* https://arxiv.org/pdf/2010.10397.pdf.
91. Chandrasekharan, E. et al. (2017). You Can't Stay Here: The Efficacy of Raddit's 2015 Ban Examined Through Hate Speech. *Proceedings of the ACM on Human-Computer Interaction,* Vol. 1, Article 31, pp. 1–22. https://doi.org/10.1145/3134666.
92. Aliapoulious, M. et al. (2021). An early look at the Parler online social network. *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 15, pp. 943–951. https://arxiv.org/pdf/2101.03820.pdf.
93. Dickson, E. J. (2021). The Qanon Community Is in Crisis—But On Telegram, It's Also Growing. *Rolling Stone Magazine.* https://www.rollingstone.com/culture/culture-news/qanon-telegram-channels-increase-1117869/.
94. Nickas, J., Isaac, M., & Frenkel, S. (2021). Millions Flock to Telegram and Signal as Fear Grows Over Big Tech. *The New York Times.* https://www.nytimes.com/2021/01/13/technology/telegram-signal-apps-big-tech.html.
95. Johnson, N. F. et al. (2019). Hidden resilience and adaptive dynamics of the global online hate ecology. *Nature, 573,* 261–265. https://www.nature.com/articles/s41586-019-1494-7.
96. Menn, J., Li, K., & Culliford, E. (2021). Parler partially reappears with support from Russian technology firm. *Reuters Internet News.* https://www.reuters.com/article/us-usa-trump-parler-russia/parler-partially-reappears-with-support-from-russian-technology-firm-idINKBN29N23N.
97. Holt, J., Brookie, G., & Brooking, E. (2021). Fast Thinking: How the Capitol riot was coordinated online. *The Atlantic Council.* https://www.atlanticcouncil.org/content-series/fastthinking/fast-thinking-how-the-capitol-riot-was-coordinated-online/.
98. Saldiva, G., Van Sant, S., & Bowman, E. (2018). Suspect Charged With 29 Federal Counts in Pittsburgh Synagogue Massacre. *National Public Radio.* https://www.npr.org/2018/10/27/661347236/multiple-casualties-in-shooting-near-pittsburgh-synagogue.
99. Greenburg, A. (2021). Far-right platform gab has been hacked—including private data. *Wired Magazine.* https://www.wired.com/story/gab-hack-data-breach-ddosecrets/.
100. Ali, S. et al. (2021). Understanding the Effect of Deplatforming on Social Networks. WebSci '12: 13th ACM Web Science Conference, pp. 187–195. https://seclab.bu.edu/people/gianluca/papers/deplatforming-websci2021.pdf.
101. Goodfellow, I. et al. (2014). Generative Adversarial Nets. Proceedings of the 27th *International Conference on Neural Information Processing Systems*, Vol. 2, pp. 2672–2680. https://arxiv.org/pdf/1406.2661.pdf.
102. Cole, S. (2017). AI-assisted fake porn is here. *Vice Magazine.* https://www.vice.com/en/article/gydydm/gal-gadot-fake-ai-porn.

103. Citron, D. (2019). The National Security Challenge of Artifical Intelligence, Manipulated Media, and "Deep Fakes." *Prepared Written Testimony and Statement for the Record. House Permanent Select Committee on Intelligence.* https://intelligence.house.gov/uploadedfiles/citron_testimony_for_house_committee_on_deep_fakes.pdf.
104. Rubio, M., & Warner, M. (2019). Rubio, warner express concern over growing threat posed by Deepfakes. *Joint Statement from Senators Marco Rubio and Mark Warner.* https://www.rubio.senate.gov/public/index.cfm/2019/10/rubio-warner-express-concern-over-growing-threat-posed-by-deepfakes.
105. Watts, C. (2019). The national security challenge of Artifical INTELLIGENCE, MANIPULATED MEDIA, and "Deep Fakes." *Prepared Written Testimony and Statement for the Record. House Permanent Select Committee on Intelligence.* https://intelligence.house.gov/uploadedfiles/clint_watts_-_house_select_committee_on_intelligence_-_ai__deep_fakes_-_13_june_2019.pdf.
106. Doermann, D. (2019). The national security challenge of Artifical intelligence, manipulated media, and "Deep Fakes." *Prepared Written Testimony and Statement for the Record. House Permanent Select Committee on Intelligence.* https://intelligence.house.gov/uploadedfiles/doermann-statement-final.pdf.
107. Clark, J. (2019). The national security challenge of Artifical intelligence, manipulated media, and "Deep Fakes." *Prepared Written Testimony and Statement for the Record. House Permanent Select Committee on Intelligence.* https://intelligence.house.gov/uploadedfiles/clark_deepfakes_sfr.pdf.
108. Deepfakes. (2019). FaceSwap Manifesto. *Github pages.* https://github.com/deepfakes/faceswap#manifesto.
109. Radek. (2018). Myfakeapp. *Bitbucket.* https://bitbucket.org/radeksissues/myfakeapp/src/master/.
110. Boyd. M. (2021). Deepfakes and LinkedIn: malign interference campaigns. *Malwarebytes Labs, Social Engineering.* https://blog.malwarebytes.com/social-engineering/2019/11/deepfakes-and-linkedin-malign-interference-campaigns/.
111. Collins, B. (2019). Facebook says a pro-Trump media outlet used artificial intelligence to create fake people and push conspiracies. *National Broadcasting Company News.* https://www.nbcnews.com/tech/tech-news/facebook-says-pro-trump-media-outlet-used-artificial-intelligence-create-n1105951.
112. Associated Press. (2021). Cheerleader's mom accused of making 'deepfakes' of rivals. The Associated Press https://apnews.com/article/pennsylvania-doylestown-cheerleading-0953a60ab3e3452b87753e81e0e77d7f.
113. Kerras, T., Laine, S., & Aila T. (2019). A style-based generator architecture for generative adversarial networks. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).* https://arxiv.org/abs/1812.04948.
114. nVidia Labs. (2019). Style-Gan Repository. *Github.* https://github.com/NVlabs/stylegan.
115. Kerras, T. et al. (2020). Analyzing and improving the image quality of StyleGAN. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).* https://arxiv.org/abs/1912.04958.
116. nVidia Labs. (2020). Style-Gan2 Repository. *Github.* https://github.com/NVlabs/stylegan2.
117. Zakhorav, E. et al. (2019) Few-shot adversarial learning of realistic neural talking head models. *2019 IEEE/CVF International Conference on Computer Vision (ICCV).* https://arxiv.org/pdf/1905.08233.pdf.
118. Gosse, C., & Burkell, J. (2020). Politics and porn: How news media characterizes problems presented by deepfakes. *Critical Studies in Media Communications,* 30(5). https://doi.org/10.1080/15295036.2020.1832697?journalCode=rcsm20#.YPrydo5KhPa.
119. Littell, J. (2019). Don't Believe Your Eyes (or Ears): The Weaponizations of artificial intelligence, machine learning, and Deepfakes. *War on the Rocks.* https://warontherocks.com/2019/10/dont-believe-your-eyes-or-ears-the-weaponization-of-artificial-intelligence-machine-learning-and-deepfakes/.

120. Weiss, M. (2019). The hero who betrayed his country. *The Atlantic.* https://www.theatlantic.com/international/archive/2019/06/estonia-russia-deniss-metsavas-spy/592417/.
121. Nemtsova, A. (2019). What to Do When the Russian Government Wants to Blackmail You. *The Atlantic.* https://www.theatlantic.com/international/archive/2019/03/russia-government-blackmail-kompromat/585850/.
122. Tung, L. (2016). Adobe's VoCo voice project: Now you really can put words in someone else's mouth. *ZDnet.* https://www.zdnet.com/article/adobes-voco-voice-project-now-you-really-can-put-words-in-someone-elses-mouth/.
123. Dessa Corporation. (2019). RealTalk: We Recreated Joe Rogan's Voice Using Artificial Intelligence. YouTube. https://youtu.be/DWK_iYBl8cA.
124. Kosoff, M. (2017). How Russia secretly orchestrated dozens of U.S. protests. *Vanity Fair.* https://www.vanityfair.com/news/2017/10/how-russia-secretly-orchestrated-dozens-of-us-protests.
125. Kawthekar, P., Rewari, R., & Bhooshan, S. (2017). Evaluating generative models for text generation. *Stanford University.* https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1174/reports/2737434.pdf.
126. Brown, T. et al. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems, 33.* https://arxiv.org/abs/2005.14165.
127. Anand, A. (2020). Another 10 gems of GPT-3. *Dev.to Blog.* https://dev.to/amananandrai/another-10-gems-of-gpt-3-2639.
128. Huntington, H. (2017). The affects and effect of internet memes: assessing perceptions and influence of online user-Generated Poltical discourse as media. *Colorado State University Department of Journalism and Media Communication.* https://mountainscholar.org/bitstream/handle/10217/183936/Huntington_colostate_0053A_14303.pdf.
129. Hutson, M. Robo-writers: the rise and risks of language generating AI. *Nature News Feature.* https://www.nature.com/articles/d41586-021-00530-0.
130. Xiao, H et al. (2015). Is feature selection secure against training data poisoning. *Proceedings in the 3nd International Conference on Machine Learning.* http://proceedings.mlr.press/v37/xiao15.pdf.
131. Xiao, H. et al. (2015). Support vector machines under adversarial label contamination. *Neurocomputing, 160*, 53–62. https://www.sec.in.tum.de/i20/publications/support-vector-machines-under-adversarial-label-contamination/file/main-revision.pdf.
132. Nguyen, T. et al. (2019). Deep learning for Deepfakes creation and detection: a survey. *Cornell University Arxiv.* https://arxiv.org/pdf/1909.11573.pdf.
133. Acemoglu, D., & Robinson, J. (2019). The narrow corridor. states, societies, and the fate of liberty. Penguin Random House Publishing.

Joe Littell enlisted in the Army in 2003 as an infantryman and attained the rank of Sergeant before commissioning in 2010. Upon commission, Major Littell has served as a Platoon Leader, Company Executive Officer, and Battalion Logistics Officer while assigned to the 83rd Chemical Battalion. As a 1LT, MAJ Littell applied for, assessed, and completed the Psychological Operations Qualification Course and served within the ARSOF community as a Tactical Detachment Commander and Company Commander with 9th PSYOP Battalion (Airborne). MAJ Littell currently serves as a research scientist at the Army Cyber Institute at West Point on the Information Warfare team working on computational propaganda, narrative warfare, radicalization, and microtargeting through publicly and commercially available data. He holds a BS in Computer Science from the University of South Florida and a MS in Data Science from Duke University.