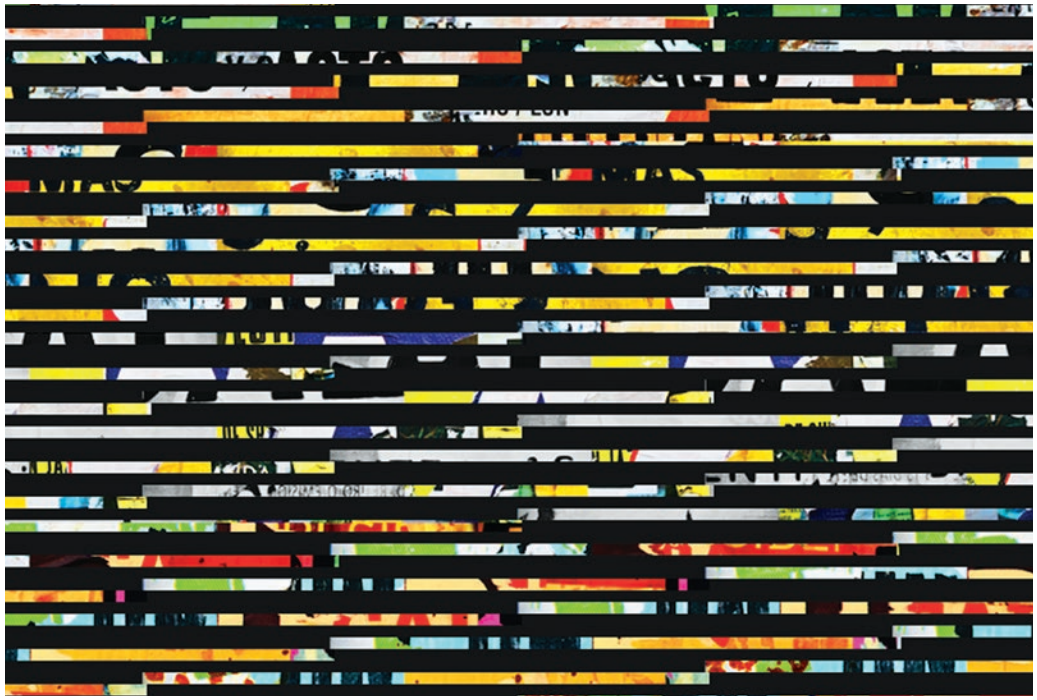
A New Wave of Censorship:

Distributed Attacks on Expression and Press Freedom

DANIEL ARNAUDO

May 2018



CIMA
CENTER FOR INTERNATIONAL MEDIA ASSISTANCE

NED
National Endowment for Democracy
Supporting freedom around the world

CIMA
CENTER FOR INTERNATIONAL MEDIA ASSISTANCE

ABOUT CIMA

The **Center for International Media Assistance (CIMA),** at the National Endowment for Democracy, works to strengthen the support, raise the visibility, and improve the effectiveness of independent media development throughout the world. The center provides information, builds networks, conducts research, and highlights the indispensable role independent media play in the creation and development of sustainable democracies. An important aspect of CIMA's work is to research ways to attract additional US private sector interest in and support for international media development.

CIMA convenes working groups, discussions, and panels on a variety of topics in the field of media development and assistance. The center also issues reports and recommendations based on working group discussions and other investigations. These reports aim to provide policymakers, as well as donors and practitioners, with ideas for bolstering the effectiveness of media assistance.

Center for International Media Assistance
National Endowment for Democracy

1025 F STREET, N.W., 8TH FLOOR
WASHINGTON, DC 20004
PHONE: (202) 378-9700
FAX: (202) 378-9407
EMAIL: CIMA@ned.org
URL: https://cima.ned.org

Mark Nelson
SENIOR DIRECTOR

Daniel O'Maley
PUBLICATION EDITOR

National Endowment for Democracy
Supporting freedom around the world

A New Wave of Censorship: Distributed Attacks on Expression and Press Freedom

MAY 2018

Contents

ABOUT THE AUTHOR

Daniel Arnaudo is a Senior Program Manager in the governance department of the National Democratic Institute (NDI) in Washington, DC. In this capacity, he covers the intersection of democracy and technology with a special responsibility to develop projects countering and tracking disinformation worldwide. Concurrently, he is a Research Fellow with the Igarapé Institute of Rio de Janeiro and a Cybersecurity Fellow at the University of Washington's Jackson School of International Studies where he has worked on projects in Brazil, Myanmar, and the United States. Recently, he has also collaborated with the Oxford Internet Institute's research group on Computational Propaganda. His research focuses on online political campaigns, digital rights, cybersecurity, as well as information and media literacy. He earned masters degrees in Information Management and International Studies at the University of Washington by completing a thesis on Brazil and its Internet Bill of Rights, the *Marco Civil da Internet*. In the past, he has worked for the Arms Control Association, the Carnegie Endowment for International Peace, and the Carter Center. He has also consulted for a wide range of organizations including Microsoft, the Center on International Cooperation at New York University, and NASA.

The views and opinions expressed in this report are those of the author.

New Forms of Censorship by Distributed Attacks on Expression and Press Freedom

As soon as the internet became an important tool for sharing independent news and empowering citizens to speak their minds, authoritarian governments and their political allies started to seek ways to censor and block content that might undermine their grip on power. Initially, this entailed censoring the content of individual pages or users, blocking websites, and at times even cutting off internet access to entire communities, cities, and countries.

This represented a direct form of censorship in which regimes suppressed objectionable content by removing it from the public sphere. Now, however, authoritarian actors are becoming more sophisticated in the strategies they use to curtail access to information and freedom of the press. They have developed novel, distributed forms of censorship that utilize new technologies, such as automated social media accounts and selective throttling of bandwidth, to constrain news circulation and the public discourse. While these new tactics are often less perceptible to the general public, they have the overall impact of fundamentally undermining an open and independent news media ecosystem that is the bedrock of democracies.

While these new tactics are often less perceptible by the general public, they have the overall impact of fundamentally undermining an open and independent news media ecosystem that is the bedrock of democracies.

Whereas traditional forms of censorship seek to overtly block content from circulating, these new forms of censorship are less focused on totally removing content from the public sphere. Rather, they seek to disrupt the media ecosystem by alternatively overwhelming it with content, often hyper-partisan stories and even downright disinformation, or by chilling communication through slower internet speeds and self-censorship induced by overt surveillance. When viewed together, these tactics represent what can be called a *distributed attack on expression and press freedom*. Just as a distributed denial of service (DDoS) attack renders a website inaccessible through a flood incoming traffic from many different sources—requests that overwhelm the server and make it impossible to operate—these new threats combine to overwhelm public institutions, the media, and the democratic principles that undergird civil society.

As in a DDoS assault, in isolation, none of the incoming requests is out-of-the-ordinary or seen as malicious, but together the effect is paralyzing. And because the attacks are distributed—they are launched from different systems operating in conjunction with each other rather

© iStock by Getty Images / LorenzoPatoia

Though sowing doubt and cynicism through false narratives is not a new tactic, it is one that has taken on new manifestations in the age of social media. Little attention, however, has been given to the chilling or stifling effects of disinformation campaigns on freedom of expression and the press.

than a single source—they can also be more difficult to address. In essence, authoritarian regimes and other political actors interested in manipulating the public sphere have utilized new mechanisms to influence the flow of information in ways that undermine the ability of journalists to report stories, disseminate their content, and also the capacity of citizens to assess what information is reliable and accurate. These conditions curtail the ability of media to fulfill their role to inform the public, stifle discourse in civil society, and have the longer-term effect of eroding public trust in the news media.

Over the past couple of years there has been a heightened, global awareness about the impact that disinformation campaigns can have on political processes, both in authoritarian countries and democracies. Though sowing doubt and cynicism through false narratives is not a new tactic, it is one that has taken on new manifestations in the age of social media. Little attention, however, has been given to the chilling or stifling effects of disinformation campaigns on freedom of expression and the press. Even when audiences remain unconvinced by disinformation or propaganda, the distribution of intentionally misleading information or false accusations can still achieve its goal of undermining a free and open news media ecosystem by crowding out reliable content and re-directing the topics of public discussion. Furthermore, while disinformation campaigns may be directed by state actors or their direct proxies, in many cases they are independently amplified by real individuals acting on their own accord. This is another way in which these new mechanisms are distributed, and therefore harder to counteract. The blurred lines and multiple vectors from which this type of content emanates make addressing these distributed attacks on freedom of expression and the press incredibly complex.

Complicating matters even further, the media ecosystem has also undergone a dramatic shift over the past twenty years. New tech platforms, like Google, Facebook, and Twitter now play a central role in the global circulation of news, and have upended traditional news organizations by both altering distribution mechanisms and reconfiguring the advertising market, which has had an extremely negative impact on privately-owned news outlets traditionally supported by advertising. In terms of these new forms of censorship, the business models of social media corporations have generated perverse incentives that have exacerbated the problem at times, particularly when it comes to the circulation of dis- and misinformation. In many instances, the lack of transparency practiced by these private companies has also obfuscated the true scope of these challenges and how are they affecting societies.

Through case studies in Ukraine, Turkey, the Philippines, Bahrain, and China, this report will elucidate how new forms of distributed online censorship have undermined freedom of expression and press freedom in ways that defy conventional notions of control. While these various methods are often used in conjunction with each other, we gain a better sense of how they operate by highlighting them separately. Indeed, journalists, public and private sector entities, civil society, and others concerned with the development of news media ecosystems must understand how these techniques operate in order to construct and implement effective responses.

Unfortunately, there is no easy policy solution to the new forms of censorship because these practices often take advantage of technologies are benign or even beneficial when utilized for other means. For example, an automated social media account that warns people of traffic congestion is quite different from one that disseminates state propaganda. Ultimately, we need to develop responses that address the problem while at the same time do not undermine freedom of expression or the development of new technologies that benefit society.

While these various methods are often used in conjunction with each other, by teasing them out as much as possible we gain a better sense of how they operate, and therefore, how public and private sector entities and broader civil society can respond.

The Dictator's Digital Dilemma Reexamined

The rise of news forms of online censorship begs us to reexamine how governments negotiate news media ecosystems in an era where the internet is a primary communications tool. The growing importance of the global network as a tool to organize economic growth has meant that even authoritarian regimes—those most wary of allowing citizens access to independent news and information—have often allowed access to digital networks.

Indeed, we now see that the internet, while providing an essential communication tool to reformers and opposition groups, also provides a useful tool for dictators and their allies to surveil and censor.

In opening their economies to the world, they also afforded citizens the benefits that come with connecting to global communications networks, such as broader access to information. These benefits in turn come with a loss of control over the ideas that their citizens encounter. As a result, the leaders of such countries are faced with a "dictator's digital dilemma," determining whether the risk of opening their networks is worth the economic rewards.[1]

Certain scholars argued that digital communication technologies would be liberating,[2] opening dictatorships or struggling democracies to different sources of news and information.[3] This theory has been contested, particularly by scholars such as Evegeny Morozov, who argued that if regimes failed to confront online networks through various forms of censorship or surveillance, they would eventually resort to traditional methods of physical or legal suppression of opposition. Indeed, we now see that the internet, while providing an essential communication tool to reformers and opposition groups, also provides a useful tool for dictators and their allies to surveil and censor. Authoritarian regimes are using new technological methods to pursue their opponents online.[4]

This dilemma is evident in several Middle Eastern countries since 2011's Arab Spring, when countries from Tunisia to Egypt to Bahrain were faced with choices over continuing their censorship regimes or allowing their citizens access to new sources of information. The internet has developed rapidly in the past six years, as more people have connected through mobile devices, using fewer websites and blogs and more social networks and messaging services. The network is increasingly encrypted through HTTPS for websites, or end-to-end encryption through secure messengers such as WhatsApp and Signal. These are global trends, but particularly significant for the Middle East's democratization movements and conflicts that have developed since 2011.[5] Certain countries in the region have moved in positive directions since the events of that time,

such as Tunisia, and opened their countries' economies, networks and societies to the world while democratizing their political systems.[6] Others have re-entrenched the dictatorships that exist, as in Bahrain or Saudi Arabia, while still others that moved towards democratic politics for a period are again reverting to authoritarianism, as in the case of Egypt.[7]

Egypt is a particularly important case in the sense of internet control and censorship because it also provides an example of a country that completely shut down its networks for a time, one of the largest national internet disconnections in history.[8] Precedents were either much smaller, such as Myanmar's decision to cut connections in 2007, or more limited, as in Iran's internet slow down during disputed elections in 2009.[9] India's blockages in selected states in 2016 represent a regionalized version of this phenomenon.[10]

© Hang Dinh/Shutterstock.com

Internet shutdowns are only the bluntest instrument in the large toolkit of authoritarian or semi-authoritarian rulers, and they come at a cost.

Internet shutdowns are only the bluntest instrument in the large toolkit of authoritarian or semi-authoritarian rulers, and they come at a cost. A report from the Brookings Institution estimated that the costs of 81 internet and social network shutdowns during a period from 2015 and 2016 totaled over $2.4 billion in dictatorships ranging from Saudi Arabia to Ethiopia, as well as in democracies including Brazil and India.[11] In all cases, there were clearly economic costs to shutting down the internet, or even blocking individual pages, domains, or social networks. HTTPS makes it very difficult for regimes to block specific pages a user requests from a domain because the request is encrypted, and as a result many countries have increasingly had to block entire networks, as Turkey did for Twitter in 2015,[12] or China and Iran now do for Facebook and Twitter.[13] They are also easily perceivable; inevitably, a state will come under criticism for shutting down the internet, and not only for the costs, but also because of the political implications of blocking a key avenue of communication, information, and expression for the citizens.

Blocking tools for websites and individual pages are a simpler form of censorship, and often ineffective when access is defined by large social networks rather than email, individually hosted web pages and forums. What's more, government agencies, police, emergency services and command and control systems also rely on the same networks, both social and technical. Because of these costly trade offs, regimes are relying on new technological tools to monitor and disrupt the flow of news and information online, ranging from automated accounts, analyzing fast-expanding stores of data, or manipulating algorithms on which billions now depend for content.

The New Tools for Democratic Disruption

The new methods authoritarian governments and other undemocratic actors are using to disrupt and manipulate democratic dialogue online are designed to allow the internet to continue to operate, while providing greater control over how information is disseminated and reaches the intended audience.

Governments and other entities can make use of large stores of data that private entities have gathered on individuals to target them very precisely and filter what they see online.

More importantly, some of these techniques can be used not only domestically, but also abroad. This is feasible given that they often employ global social media platforms. These tactics include the use of networks of automated accounts, as well as individuals employed by the state or a private company with connections to the state operating multiple profiles to post messages, support others, and influence trending topics on social networks. Other forms of computational propaganda include the manipulation of algorithms to change the topics of conversation, as well as the pages, posts, advertising and other content that users see. Governments and other entities can make use of large stores of data that private entities have gathered on individuals to target them very precisely and filter what they see online.[14] This power to change the trends of social media is combined with the influence these social networks and topics can have on traditional media. For instance, if stories go viral on Facebook or Twitter, they are often picked up on television or in newspapers and online outlets. As a result, automated tactics can be amplified by various forms of media.

What are these new tools and tactics precisely, and how do they differ from the ones that came before? How are these new forms of online censorship subtler than the ones that were used earlier, and do we have tools for monitoring and tracking them so that we can explain them to others working in the media, monitors and more broadly within civil society? What are the responses that those in media, government, and civil society can formulate to encourage the proliferation of a free, open, and democratic networked public sphere?[15]

The following sections of this report contain examples of how these new forms of censorship are operating in different authoritarian and contested democratic environments all over the world. The examples illustrate techniques that are being used by regimes to monitor populations, censor citizens, block networks, and

affect social media that often originate or are hosted in democratic countries. In each case, the stifling impact on the news media is clear, from the disruption of content distribution to the de-legitimatization of sources. Evident too in each case is a grave threat to the possibility of democratic dialogue, and the opportunities this creates for resurgent authoritarianism.

The Tools and Tactics of New Forms of Censorship

Troll farms

Localized shutdowns/ slowdowns

Dedicated moderation networks for censorship

Bot networks

IP/website/ network blocking

Automated systems to filter for political content

Distributed denial of service (DDoS) attacks

Personal data exfiltration and expropriation

Algorithms to report and remove political speech and moderate content

Abuse of terms of service to block accounts and remove content on social networks

Criminalization and tracking of online political speech

Mal-information: hacking and leaking of media, civil society and other private information for political, commercial, or personal gain

State Sponsored Trolling: Russia's Efforts to Overload Ukraine's Media Ecosystem

As social media has become an increasingly important arena for the circulation of news and the formation of public opinion, marketers and political advocates of all stripes have taken to these platforms to promote their brands and messages through large groups of social network users who purposefully create content and interact with other users.

The motive is frequently not to reach individuals with information that could potentially be useful, but rather to flood the public sphere with content—often false—and to make it incredibly difficult for citizens to filter what information can be trusted and what cannot.

However, these influence campaigns take on a whole other dimension when they are funded by states with deep pockets and enacted with specific geopolitical intentions. In these instances, the motive is frequently not to reach individuals with information that could potentially be useful, but rather to flood the public sphere with content—often false—and to make it incredibly difficult for citizens to filter what information can be trusted and what cannot. This is what happened in 2014 in Ukraine when the Russian government launched a concerted effort to disrupt the news media ecosystem amid the political upheaval of the EuroMaidan protests. The use of state-sponsored trolls to spread disinformation and to attack Ukrainian journalists online was one of the primary tactics that the Russian government employed.

Russia's ability to successfully coordinate these efforts in Ukraine was based on several important factors. First, the Russian government had already developed a network of paid posters (trolls) under the government of Dmitry Medvedev. President Medvedev embraced social media and became known as the "blogger in chief" for his use of blogs and other social media during that period.[16] The first instances of organized networks of paid posters (trolls) and automated accounts (bots) connected to the government surfaced at this time. However, in contrast to the campaigns that would come to attack democracies in other countries, at the onset these activities were mostly to promote and draw attention to the president's writings and other content online.[17] When Putin resumed the presidency in 2012, he implemented a more aggressive internet policy that included more sophisticated filtering systems and the weaponization of troll networks to attack opponents. Companies such as the Internet Research Agency, a quasi-independent organization with deep connections to the state, emerged at this time. This St. Petersburg based group brings together a complex of programmers, spammers, and simple computer users who

design and participate in varying online campaigns against opposition candidates, parties, and other movements.[18] The Russian government also used these capabilities to attack political opponents more directly, for instance by hacking their emails and distributing them through these same networks, spreading malware on their computer systems, or engaging in Distributed Denial of Service (DDoS) attacks against their websites and other networks. These tactics, combined with the oppressive legal and political climate, seriously affected the way that the media operated in Russia. Those few outlets that remained independent had to develop strong cybersecurity practices to protect sources, keep communications within the newsroom private, and keep their operations online. Further, Russian websites and social networks such as Odnoklassniki, VKontakte, Yandex, mail.ru were incredibly popular in Ukraine and other post-Soviet countries. This cybernetic connection with Russia as well as the fact that many Ukrainians got their news from Russian language television and radio made the country very susceptible to a disinformation campaign.[19] Indeed, Ukraine represented one of the most fertile places in the world for a new kind of propaganda. This intervention, which began in 2014, showed the power of these networks to extend their reach outside of Russia.

The Ukrainian revolution in 2014 represented one of the Russian government's greatest fears: a large, post-Soviet country moving away from an alliance with Russia and turning towards Europe. Its 2004 Orange Revolution had resulted in the defeat of Russia's ally Viktor Yanukovych in an election that pitted him against a Western aligned, European-oriented candidate Viktor Yushchenko. Six years later, Yanukovych regained power in another election only to be challenged in 2013 by another round of protests against corruption and a decision by his government to cancel an agreement to move towards integration with the European Union. The protests coalesced in Maidan Square in Kiev and the movement they generated came to be known as the EuroMaidan.

The massive, nation-wide protest paralyzed the country. In February of 2014, as Yanukovych and many of his key ministers fled to Russia and parliament called for special elections to replace him, disguised Russian soldiers took control of government buildings and strategic infrastructure in Ukraine's Crimea region. In a referendum that November that occurred under Russian military occupation and was denounced as illegitimate by the west, Crimea was incorporated into the Russian Federation. Further Russia-backed agitations for succession have since embroiled Ukraine's Donbas region.

Ukraine represented one of the most fertile places in the world for a new kind of propaganda. This intervention, which began in 2014, showed the power of these networks to extend their reach outside of Russia.

On the whole, these tactics represented a distributed form of attack on freedom of expression and the press because they sought to hinder the ability of journalists to communicate the news and prevented Ukrainian citizens from being able to easily access high-quality information.

These events became foci of disinformation campaigns by the Russian government, from denying the assistance that Russian forces gave to paramilitaries in Ukrainian regions, to the downing of a Malaysian Airways flight by those groups during the war.[20] Research by the Oxford internet Institute's Computational Propaganda project shows that networks of "trolls" or paid social media accounts have been particularly prevalent in Ukraine throughout these events, and at relatively cheap cost. Accounts manipulated by paid users to post about specific topics or "like" other posts, accounts or pages cost as little as US $0.40 to $0.90 on social networks such as Facebook, Twitter, and VK.[21] These networks were often based in Russia or other former Soviet republics. While in some cases these trolls, often amplified by bots, spread messages based on common themes and central organizational principles, in other cases they did so in more decentralized and multifarious ways.

Once the protests began, these networks developed quickly. Research showed that various bot networks were created during both the 2013 EuroMaidan protests and the beginning of the conflict in Eastern Ukraine in 2014.[22] These bots and trolls were used to amplify content that supported the Russian narrative that the EuroMaidan movement was a Western-backed coup, attack users who objected to this narrative, confuse users about facts on the ground, or encourage various hashtags or topics to trend on social networks. Bots or trolls have even been used to monitor real users for violations of the terms of service and report them with the goal of getting them banned or suspended. In one case, a journalist had their Facebook account disconnected for posting about the downing of a Malaysian Airways commercial airliner MH17 during the war for Eastern Ukraine.[23] Bots and trolls sent thousands of requests for takedowns to Facebook and other moderation teams, which banned or blocked user accounts tied to media or others in civil society. On the whole, these tactics represented a distributed form of attack on freedom of expression and the press because they sought to hinder the ability of journalists to communicate the news and prevented Ukrainian citizens from being able to easily access high-quality information.

Ukrainians have developed some responses to these attacks on their information ecosystem. For example, at the Kyiv-Mohyla School of Journalism, a group of individuals formed the fact-checking initiative StopFake.org. This organization counters false news narratives pushed by Russia by identifying, analyzing, and discrediting over 1,000 stories on social media since its formation in 2014.[24] They also broadcast reports on the propaganda and false narratives they find, which they distribute on YouTube and Facebook. This combination of network

analysis, identification of content, dissection of propaganda, and use of video and social media provides an effective example of how media can evolve and respond to these new challenges.[25]

Ukrainian journalists have played a major role in exposing bot networks and the use of Russian computational propaganda. The news website Texty.org.ua did a comprehensive analysis of the groups that formed to counter the current Ukrainian government and published a website that included graphical examples of how the online network functioned.[26] This combination of data scientists, graphic designers, and journalists demonstrates a powerful example of how new forms of journalism—by revealing how the disinformation networks are formed and administered—can counter new forms of propaganda and censorship. This model is especially powerful when applied with traditional forms of narrative journalism.

The Ukrainian government attempted to form a user base of social media agents to counter false narratives, and registered 40,000 individuals to work to oppose false narratives. However, the government has not been able to confirm that they have used this base in any consistent way.[27] Ultimately, the Poroshenko administration chose a more blunt strategy, banning Russian television and radio from Ukrainian networks and blocking Russian social media sites such as VKontakte and Odnoklassniki, as well as the Yandex search engine.[28] It was a questionable decision, as this kind of blanket censorship seriously affects freedom of expression in democratic society, and with dubious effectiveness, given the numerous ways to pierce the ban, such as VPNs or encrypted networks.

Ukraine is currently wrought by civil strife, and unfortunately these fractures are reflected in its social networks, which have been exploited by Russia and its allies. The country provides examples of how networks, both human and robotic, can shape narratives about events and people, but also how new kinds of media organizations like StopFake and Texty can begin to describe and counter these narratives by identifying the networks that propagate them and the content they are sending. Simultaneously, they are working to push back on these narratives by explaining why the stories are wrong and also how to use social media to discredit them. As a result, Ukraine is both a sign of how new forms of distributed censorship can operate in contested contexts, and how civil society and media can begin to form effective responses.

© Mariia Golovianko / Shutterstock.com

Ukraine is currently wrought by civil strife, and unfortunately these fractures are reflected in its social networks, which have been exploited by Russia and its allies.

Domestic Trolling: Shaping the Public Dialogue in Turkey

Developments in Turkey over the past five years provide another example of how authoritarian state agencies use large networks of pro-government users to undermine the free exchange of ideas. While "troll armies" are becoming increasingly prevalent throughout the world, Turkey exemplifies how these tools are being turned on their own populations to create a new form of distributed censorship that starves citizens of reliable news and information, and makes the work of independent journalists incredibly challenging.

Turkey's democratic institutions have been severely challenged in recent years, as President Recep Tayyip Erdoğan has changed the constitution to empower the executive and significantly cracked down on press freedom.

Turkey's democratic institutions have been severely challenged in recent years, as President Recep Tayyip Erdoğan has changed the constitution to empower the executive and significantly cracked down on press freedom. His government has jailed more journalists than any other country in the world and has shuttered or threatened more than 150 media outlets in the wake of a military coup against his regime in 2016.[29] Some of these organizations and journalists have been designated as security threats, but many have been attacked for challenging the official government narrative or not giving sufficient support to the regime and criticizing the military for its role in the plot.[30] In addition to these traditional censorship measures such as shutting down news outlets or jailing journalists, Erdoğan has moved aggressively to challenge the opposition in the online space. The 2013 protests in Istanbul against the destruction of public space Gezi Park caught the attention of many citizens throughout the country, online and through social networks, and quickly became a touchstone for the opposition movement. Since then, Erdoğan's government has worked in various ways to change the narratives and shut down opposition voices. Beyond blocking pages there are four major components to these online attacks on the media, and civil society groups that oppose the government's aims:

- Attack opposition social media accounts through networks of trolls and bots. Often these coordinated attacks are complemented by the regime's supporters working in direct coordination with government agencies.
- Lodge complaints with Twitter and other social networks against accounts that are challenging the regime in hopes that the platform will pull done the content.
- Hack journalists accounts and expose their private conversations to the public.
- Prosecute journalists for news and opinion pieces they post online.

The first component comes through a network of supportive social media accounts. The central node in the network of these campaigns was often a group of over 6,000 supporters attached to the "New Turkey Digital Office" that promoted ideas supporting the regime and attacked those who did not agree with the government's perspective.[31] These networks were also capable of activating thousands of followers in online social networks to support these campaigns. Government affiliated and supporting groups increased their use of these tactics in March 2014 when they focused on defending Erdoğan and his allies from accusations of corruption that surfaced on Twitter from an account known as @ oyyokhirsiza. This account leaked confidential information that showed questionable business dealings of his Minister of Communication, Binali Yıldırım, and his son. The Shorenstein Center at Harvard defines this kind of campaign as a form of "malinformation" in that it describes information that is often true, hacked, and leaked to discredit the user as well as the ideas and objectives. Erdoğan pledged to wipe out Twitter and even temporarily blocked it. However, civil society and opposition groups responded by using VPNs and other workarounds to virtually tunnel out of the country, and spread information about the shutdown through the hashtag #TurkeyBlockedTwitter that helped end the blockage relatively quickly.[32] The use of state-led troll networks brings to bear state-sponsored campaigns combined with members of the public that are influenced by them to post their own social media. This constitutes a distributed attack on democratic discourse, through the spread of state propaganda and the diminution of opposing themes, accounts, and content.

A second tactic used by the government was to tap these same networks to attack journalists by submitting complaints against their content on Facebook and Twitter. The objective of these repeated complaints from multiple users was to encourage social media platforms to remove the content. This technique ramped up in 2014. In the first half of the year, there were roughly 200 such complaints lodged on Twitter, while this doubled to more than 400 in the second half of the same year. These trends only increased over time, as it became one of the largest supplicants of account deletion or content removal on the network through 2017.[33] After the 2016 coup attempt, attacks on journalists, their organizations, and others in civil society extended to the online sphere. User accounts associated with the regime together with supporters spurred on by a climate of hatred toward any opposition launched attacks on anyone critical of the government. Female journalists became common targets. A study of tweets attacking journalists in 2016 by the International Press Institute (IPI) found that almost 10 percent of them were sexually related comments directed overwhelmingly at women. Other methods catalogued included humiliating tweets (9 percent) intimidating content

A second tactic used by the government was to tap these same networks to attack journalists by submitting complaints against their content on Facebook and Twitter. The objective of these repeated complaints from multiple different users would encourage social media platforms to remove the content.

(10 percent) and "Threats of violence, other abusive behaviors, legal threats and technical interferences (72 percent)".[34] These networks have encouraged a climate of fear, self-censorship, and suppressed social and political expression online in various forums.

The government and its allies have also moved to attack journalists via a third vector, through hacking their private accounts and spreading their own confidential conversations with sources, coworkers, and other contacts. IPI found 20 cases of journalists having their accounts hacked in this period, usually announced by the culprits taking control of their Twitter account and posting messages supporting the regime. For instance, when the journalist Can Atakli's account was hacked the attackers scrawled, "I apologise to our honourable president to whom I was unfair and bashing all this time with my libels and insults" with a picture of the President attached; his direct messages were meanwhile shared in online forums.[35] It should be noted that these types of attacks not only impact the journalists who are the targets, but they also serve to sow doubt and confusion among the broader population about who to trust. They create insecurity as it can become more difficult to know what is real and what is false online.

These types of attacks not only impact the journalists who are the targets, but they also serve to sow doubt and confusion among the broader population about who to trust.

Finally, these tactics are combined with a fourth, more traditional tactic of simply prosecuting and jailing journalists. This is now bolstered by a new constitution that criminalizes many kinds of speech against the state or the security services. New forms of censorship, such as the use of troll armies and hackers to find incriminating materials, are more effective in combination with stringent laws against threatening state security or other equally nebulous concepts. Turkey provides a primary example of how distributed attacks on freedom of expression and the press can work in a country struggling to maintain a semblance of a democratic system. These armies of user accounts can be used in various ways: to attack opposition, identify accounts for removal under terms of service, or simply to promote the policies of the state. It is a powerful new tool in the arsenal of censorship that states can now employ, and combined with older methods, can be a force multiplier in terms of policies and ideas, encouraging a public sphere defined by the narrative of the regime, and disparaging and inciting fear in any opposition. The combination of legal, physical, and online threats has taken a toll and promoted a kind hybrid censorship that has been effective in silencing the media, confusing users, and blunting the effects of critical press.

Automated Bot Networks: Filipino Bots and the Social News Network Response

Automation brings another level of coordination and computing power to bear through distributed forms of censorship. Networks of automated accounts or bots, known as botnets, can be used to promote content, create trending topics, or attack others, generally for a relatively low investment, even compared to trolls, as individual users can operate thousands of individual accounts or even enable them to operate autonomously.[36] Though the Philippines is now more commonly invoked in the study of how media freedoms and democracy can be unwound, the country's experience also illustrates how a strong response from media organizations can push back against new forms of censorship and control.

Since the election of Rodrigo Duterte, the government has begun a campaign to eliminate drug usage in the country through harsh tactics that include mass incarceration and even vigilantism against drug dealers and users. This has led to rising attacks on people associated with the drug trade, but has also increased attacks on opposition parties, civil society, and the media. As in other contexts, these attacks have been bolstered by an increasing climate of intolerance online.[37]

Bots are especially good at inflating the importance of topics, repeating hashtags or other trends and content online, a tactic that is especially critical during elections, debates, and other moments of acute political importance. Four days after Duterte declared his candidacy, observers found examples of suspicious increases in the tags associated with his campaign rising to over 10 times the combined mentions of his rivals, likely caused by bots posting hundreds of times per minute.[38] The Philippines provides an example of how these automated systems work, but also how they can be identified and confronted via new independent media networks.

In the Philippines, the "social news network" known as Rappler has created an organization of journalists, data scientists, and ordinary users to track political campaigns that use trolls, fake "sock puppet" accounts, botnets, and other forms of manipulation to stir up and direct fervent supporter groups.[39] As in many other developing countries with less infrastructure, expensive mobile data, and less access to full-size computers or tablets, Facebook has become a particularly significant network for millions of people, often connecting through zero-rated services such as Facebook Free Basics, which provides low-income Filipinos with subsidized access to a bare-bones version of Facebook. They have uncovered a botnet

Bots are especially good at inflating the importance of topics, repeating hashtags or other trends and content online, a tactic that is especially critical during elections, debates, and other moments of acute political importance.

The the "social news network" known as Rappler has created an organization of journalists, data scientists, and ordinary users to track political campaigns that use trolls, fake "sock puppet" accounts, botnets, and other manipulation to stir up and direct fervent supporter groups.

supporting President Duterte and his party, often connected to influencers such as the former sex blogger and singer Mocha Uson.[40] Now an Assistant Secretary in the Government's Presidential Communications Operations Office (PCOO), she has repeatedly attacked opponents of the regime on social media and promoted accounts that are supportive of the government. Through a combination of data science and old-fashioned reporting, Rappler has demonstrated how Uson's popularity can direct her large follower base, and even influence the algorithm that ranks the content networks that her followers view.[41]

The organization has also profiled the use of bots by supporters of Duterte's party and campaign, and how this led to a surge in support during his election in 2016. They interviewed members of the campaign apparatus as well as organizations and companies that supported them, augmenting their reporting of the content of the messages with network analysis and interviews. These tactics paint a compelling picture of the state of the online space in the Philippines and have angered supporters like Uson to the point that she has requested that they be reclassified as a social networking group rather than a news organization.[42] Notably, this reclassification would make Rappler more accountable to Uson's office. It has also been attacked through legal means, as the government has challenged its tax status by questioning its foreign funding, and others have sued it for libel under a 2012 cybercrime law.[43] Besides the popularity of its content, the fact that the government is attempting to define Rappler as a social media company while pursuing it for tax evasion, suggests that its methods of finding and identifying government accounts while promoting opposing views have achieved a qualified but notable level of success.

© J Gerard Seguia/ZUMA Wire/Alamy Live News

Throttling Discourse: The Stifled Arab Spring in Bahrain

Governments have often engaged in forms of censorship that incorporate blocking websites, networks, or even the entire internet to control public discourse in different forms. However, website blocking can often be circumvented by technology such as VPNs that tunnel into other networks and hide the user's origin. Blocking also tends to draw public attention and outrage, as was the case in Turkey when the government blocked Twitter.

Throttling the internet—slowing the speed of user's access—provides another form of censorship that is more difficult for users to detect, to the point that they may believe their device or network has another technical issue unrelated to any form of government involvement. It is a distributed attack that covers many users who are reliant upon cellular networks to connect to their allies, friends, and family, coordinate, and generally understand social and political systems.

In the wake of the Arab Spring, several regimes in the region developed new systems for the control of their domestic internet, and Bahrain provides an important example. As a small gulf kingdom under the control of a single family, the regime often censors speech that is harmful to its image, whether political, social, or related to security issues. The media in the country is tightly controlled; only outlets that are friendly to the government can operate, and multiple journalists have been jailed for covering taboo topics. Television stations have also been closed for similar reasons.[44] Because of this restricted media environment, the internet provides a key conduit for citizens to access information about the world. Besides documented cases of blocking, the country engages in widespread surveillance of activists and other opponents of the regime, including with software that hacks the phones, computers, and other devices.[45] These advanced surveillance systems are marketed by corporations as a method for law enforcement or intelligence investigations, but in the hands of authoritarian regimes can also be used to stifle opposition, track dissidents, incite fear in citizens and inhibit the ability of activists and journalists to cultivate sources or work within teams. Cybersecurity thus becomes a critical element of operational security for any media organization working in these contexts.

Throttling the internet—slowing the speed of user's access—provides another form of censorship that is more difficult for users to detect.

This throttling is a new kind of technique because it does not completely shut off access, but slows it and makes it difficult for groups or individual users to coordinate and share information as it is happening in real time.

In 2016, protests over the revoking of the citizenship of a popular cleric around the town of Duraz drew national attention. The regime responded by limiting the speed of different mobile services, and severing 3G and 4G connectivity, essentially rendering access to a slower, more basic velocity well below broadband, which is also much more difficult to encrypt and transmit through modern applications.[46] This mirrors activities that occurred in Iran, where users were not cut off from access but it was significantly limited.[47] This includes in terms of their access to independent information about the state of the government, the opposition, and basic facts about their political system and society, and makes it much more difficult for them to trust in or even find free media.

This throttling is a new kind of technique because it does not completely shut off access, but slows it and makes it difficult for groups or individual users to coordinate and share information as it is happening in real time. The technique hinders the ability to organize a protest or promote opposition media, and has the benefit of masking the nature of the problem. Users may potentially think there is another kind of technical difficulty with their device, or with those they are communicating with, rather than a complete disconnection. This fits with a pattern of attacks that are no longer in the open, but rather obfuscated, and as in other authoritarian or semi-democratic states, bolstered with an increasing number of supporters entering the online space to defend the regime.[48]

Strategic Distraction and Social Surveillance: China's New Tactics to Constrain News and Information

It is well known that Chinese government has developed massive technical means to directly censor and filter information online. Indeed, the Chinese government has developed a model for the rest of the world in terms of network blocking through what has become known as the Great Firewall. This system allows the government to block access to news websites.

China's censorship is augmented by an omnipresent social media monitoring apparatus called the Golden Shield.[49] Both systems are now increasingly empowered by an army of monitors, and intelligent filtering algorithms have become extremely effective in managing content on Chinese networks. Given that the media that operate in the country are already required to obtain a license from the government and are heavily restricted in terms of the type of stories they can cover, this censorship makes China one of the most restricted environments for press freedom around the world. However, what is less well known about China's efforts to manage the information ecosystem is how it is now employing distributed forms of censorship to both strategically distract the public from contentious issues as well as employing new forms of "social credit" that provoke individual internet users to monitor others and censor themselves. These new tactics represent a fundamental threat to press freedom and access to information, and because they are more distributed and hidden, they are even more difficult to counteract.

Given that the media that operate in the country are already required to obtain a license from the government and are heavily restricted in terms of the type of stories they can cover, this censorship makes China one of the most restricted environments for press freedom around the world.

China's technological prowess as well as the size of its market give it significant leverage is setting the ground terms for tech companies to operate in the country. All domestic internet or social networking companies, such as Baidu, WeiBo and WeChat, have systems in place to register user IP addresses as well as real names and other identifying information. They participate in the Golden Shield system to proactively take down content related to sensitive subjects such as 1989 Tiananmen Square massacre or general democratic political reform. A study of various social networks and internet forums in 2013 found pervasive and rapid censorship throughout, with censors often deleting illicit content within a day.[50] Foreign companies that wish to operate in China must agree to some form of these rules, or risk being blocked by the Great Firewall while their servers are located outside of the Chinese national

internet. Facebook has been actively discussing a similar system of active censorship with the government as it has been attempting to negotiate access to the Chinese market for the past several years.[51]

Interestingly, researchers have found a relatively low level of bot activity in China.[52] Automation, however, increasingly plays a role in the form of systems that are dedicated to understanding what users are saying and taking down content automatically. Such machine learning techniques will only sharpen and augment the regime's ability to track users and take down content in real time going forward in future.[53] Intelligent systems that can identify patterns of communications, track themes, and respond to them in real time are likely to replace the army of bureaucrats, online censors and collaborative party members that currently make up the online censorship system that exists in China today. Unfortunately, such an automated system has the potential to be much more powerful and far reaching than that which exists today. In a way, this increasingly hybrid censorship system mirrors those developed by other authoritarian regimes working with bots and trolls, in that the censorship is evolving to include both human and automated elements. Bot accounts do not perform the censoring as they do in Russia or other contexts, and far fewer are found operating in a political context in China,[54] but automated systems are performing a role by blocking certain users, content, and themes across networks. The system collectively acts as an automated gatekeeper through algorithmic manipulation and other tactics, which has the effect of modifying public discourse based on the regime's priorities.

© pcruciatti / Shutterstock

The system collectively acts as an automated gatekeeper through algorithmic manipulation and other tactics, which has the effect of modifying public discourse based on the regime's priorities.

The Chinese have also become adept at generating their own content through organized teams that control their own accounts and shape discussions. Researchers from Harvard have estimated users associated with the so-called "50 cent groups" that spread government supporting narratives generate 448 million comments a year on average.[55] They conclude the goal is often to dilute the discussions of political topics, and create "strategic distraction."

The government is implementing a new Social Credit System (SCS) with the help of Chinese internet companies that may prove to push users to self-censor, and avoid sensitive subjects for fear of negative ratings that translate into a lack of privileges and access to services throughout society, basic rights. The SCS rates users and assigns scores based on such factors as their social media usage, network of friends, credit history, and shopping habits. These systems are being tested by Chinese affiliates of the online conglomerates Alibaba and Tencent. These companies are encouraging users to opt-in to these systems to gain credit bonuses and special services, but they will become mandatory for

all Chinese citizens in 2020.[56] This system has the potential to amplify self-censorship in powerful ways, as users restrict their writings, videos and other content to avoid negative commentary and score. A user is rated on characteristics, including the content they post, the number of times they have been censored or reprimanded online, and the circle of connections or "friends" they maintain. Depending on the nature of their network, even association with people with lower scores could have a negative effect on their own.

Similarly, the kind of media they are able to access independently has an effect on their views of the regime, its propaganda, and its supporters. A recent study by two Stanford University scholars found that when given the ability to access foreign news, very few younger Chinese students took the opportunity, suggesting that various forms of social and technical censorship have become deeply internalized.[57] However the research also noted that when given encouragement as well as access, the students not only consumed more foreign sources of news, but also spread it to their peers, questioned government narratives, and even sought out more external sources of information after the study had ended. Conversely, the development of the social credit system may prove to only further internalize these beliefs and practices, and the avoidance of controversial themes, users, sources and media.

Chinese methods have been replicated in several regimes in its orbit, including Vietnam, Thailand, Malaysia and Indonesia, which have all erected various forms of blocking: technical, political, and social.[58] They are also becoming a strong model for countries beyond the region such as Iran, who have also adopted a domestically bounded network, and attempted to build national social networks and services while blocking global ones such as Facebook and Twitter.[59]

The Chinese model of censorship is nurturing new forms of control that are much more difficult to confront directly. To date, the development of circumvention tools that allow internet users in China to evade the Great Firewall and gain access to content on the global internet has been an essential form of combating censorship. Technical tools such as VPNs to evade firewalls, are in many ways simpler to apply than long-term education about the importance of a free media and open access, freedom of expression, and other democratic values. Both types of education become important in countries where this controlled model is applied.

The Chinese model of censorship is nurturing new forms of control that are much more difficult to confront directly. To date, the development of circumvention tools that allow internet users in China to evade the Great Firewall and gain access to content on the global internet has been an essential form of combating censorship.

Conclusion: Toward a Collective Response

The cases examined here show indisputable trends of hidden, distributed forms of censorship around the world. They are subtler than internet shutdowns and domain blockages, although they are often deployed in tandem.

These techniques range from the documented Russian campaign to spread disinformation and interfere in political systems globally, to selective throttling in Bahrain, to armies of trolls and bots deployed in contexts such as the Philippines, Turkey and Ukraine. In China and its imitators, nationally delimited networks combined with powerful automation, big data, and monitoring systems are creating ways to replicate territorial censorship concepts globally.

These examples defy expectations that the internet would become a medium for breaking down levers of government control. Increasingly sophisticated systems will amplify these techniques, as the Chinese model shows how intelligent systems can predict and respond to individuals in increasingly rapid, effective fashion, better informed by large automated systems.[60] It is understandable why this highly regimented and regulated authoritarian society is investing so much in technology that will enable it to closely manage the growing Chinese internet.[61]

However, this study also highlights the growing responses to these threats. In Ukraine, there are several groups working to combat Russian and domestic threats to the information space, such as from StopFake, which brings together students, faculty, and alumni of the Kyiv-Mohyla School of Journalism to identify and counter false stories online. Media groups such as Texty.ua in Ukraine and Rappler in the Philippines show how journalists can partner with data scientists, graphic designers, and activists to identify fake patterns in social networks, as well as individual accounts. These networks are trackable, but will require new partnerships across social science and technical fields. Such partnerships will become increasingly valuable in confronting disinformation promoted by authoritarian regimes and their supporters, particularly to identify sources and networks quickly to respond to these trends in real time. Technology companies are developing various programs to partner with news organizations, notably Facebook's Journalism Project and Google's News Lab, and these too should respond to the censoring effects of these techniques.

Journalists have always needed the lawyers and watchdog groups to shield them from abuse and harassment and to defend their rights. What the examples in this report illustrate is that journalists need the expertise of an entirely new array of actors to protect them: data scientists, digital security experts, and digital platforms, among them. Journalists, however, may also have to play a more proactive role in conjunction with these actors. To truly neutralize these new distributed forms of censorship, civil society, the media, governments committed to democratic principles and the private sector will need to respond collectively, in a similarly distributed fashion. The internet, we now recognize, can be a tool for either the oppressor or the oppressed, but with this recognition comes an understanding that intelligent and coordinated responses can shape the existing socio-political reality, online and off.

Endnotes

1 Kerr, Jaclyn. "The Digital Dictator's Dilemma: Internet Regulation and Political Control in Non-Democratic States." Stanford, 2014. http://cisac.fsi.stanford.edu/sites/default/files/kerr_-_cisac_seminar_-_oct_2014_-_digital_dictators_dilemma.pdf.

2 Diamond, Larry. "Liberation Technology." *Journal of Democracy* 21, no. 3 (July 14, 2010): 69–83. doi:10.1353/jod.0.0190.

3 Meier, Patrick Philippe. "Do 'Liberation Technologies' Change the Balance of Power between Repressive States and Civil Society?" Fletcher School of Law and Diplomacy (Tufts University), 2012.

4 Liberation Technology: Whither Internet Control?" *Journal of Democracy* 22, no. 2 (April 2011). http://www.journalofdemocracy.org/article/liberation-technology-whither-internet-control.

5 Faris, Robert, John Kelly, Helmi Noman, and Dalia Othman. "Structure and Discourse: Mapping the Networked Public Sphere in the Arab Region," 2016. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2744915.

6 Ghannouchi, Rached. "From Political Islam to Muslim Democracy: The Ennahda Party and the Future of Tunisia." Foreign Affairs 95 (2016): 58.

7 Hessler, Peter. "Egypt's Failed Revolution." *The New Yorker,* December 26, 2016. https://www.newyorker.com/magazine/2017/01/02/egypts-failed-revolution.

8 Arnaudo, Daniel, Aaron Alva, Phillip Wood, and Jan Whittington. "Political and Economic Implications of Authoritarian Control of the Internet." In Critical Infrastructure Protection VII, 3–19. IFIP Advances in Information and Communication Technology. Springer, Berlin, Heidelberg, 2013. https://doi.org/10.1007/978-3-642-45330-4_1.

9 Howard, Philip N., Sheetal D. Agarwal, and Muzammil M. Hussain. "When Do States Disconnect Their Digital Networks? Regime Responses to the Political Uses of Social Media." *The Communication Review* 14, no. 3 (2011): 216–232.

10 Gupta, Apar, and Raman Jit Singh Chima. "The Cost of Internet Shutdowns." *The Indian Express*, October 26, 2016. http://indianexpress.com/article/opinion/columns/internet-access-government-restriction-shutdown-3102734/.

11 West, Darrell M. "Internet Shutdowns Cost Countries $2.4 Billion Last Year." Washington D.C.: Brookings, October 6, 2016. https://www.brookings.edu/research/internet-shutdowns-cost-countries-2-4-billion-last-year/.

12 Efe Kerem Sozeri. "Uncovering the Accounts That Trigger Turkey's War on Twitter." The Daily Dot, January 31, 2015. https://www.dailydot.com/layer8/twitter-transparency-report-turkey-censorship/.

13 Clark, Justin, Rob Faris, Ryan Morrison-Westphal, Helmi Noman, Casey Tilton, and Jonathan Zittrain. "The Shifting Landscape of Global Internet Censorship." Internet Monitor. Harvard Berkman Center for Internet and Society, June 29, 2017. https://thenetmonitor.org/research/2017-global-internet-censorship.

14 Krogerus, Hannes Grassegger & Mikael. "The Data That Turned the World Upside Down." *Motherboard*, January 28, 2017. https://motherboard.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win.

15 Benkler, Yochai. The Wealth of Networks: How Social Production Transforms Markets and Freedom. Yale University Press, 2006.

16 West, Darrell M. "President Dmitry Medvedev: Russia's Blogger-in-Chief." *Brookings,* November 30, 2001. https://www.brookings.edu/opinions/president-dmitry-medvedev-russias-blogger-in-chief/.

17 Kelly, John, Vladimir Barash, Karina Alexanyan, Bruce Etling, Robert Faris, Urs Gasser, and John G. Palfrey. "Mapping Russian Twitter." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, March 23, 2012. https://papers.ssrn.com/abstract=2028158.

18 Chen, Adrian. "The Agency." *The New York Times,* June 2, 2015, sec. Magazine. https://www.nytimes.com/2015/06/07/magazine/the-agency.html.

19 Ennis, Stephen. "Ukraine Hits Back at Russian TV Onslaught." BBC News, March 12, 2014, sec. Europe. http://www.bbc.com/news/world-europe-26546083.

20 Zhdanova et al.

21 Zhdanova et al.

22 Alexander, Lawrence. "The Curious Chronology of Russian Twitter Bots." *Global Voices,* April 27, 2015. https://globalvoices.org/2015/04/27/the-curious-chronology-of-russian-twitter-bots/.

23 Zhdanova et al.

24 Kramer, Andrew E. "To Battle Fake News, Ukrainian Show Features Nothing but Lies." *The New York Times,* February 26, 2017, sec. Europe. https://www.nytimes.com/2017/02/26/world/europe/ukraine-kiev-fake-news.html.

25 Kramer, Andrew E. "To Battle Fake News, Ukrainian Show Features Nothing but Lies." *The New York Times,* February 26, 2017, sec. Europe. https://www.nytimes.com/2017/02/26/world/europe/ukraine-kiev-fake-news.html.

26 Romanenko, Nadiya, Iaryna Mykhyalyshyn, Pavlo Solodko, and Orest Zog. "The Troll Network." *ТЕКСТИ.ORG.UA*, October 4, 2016. http://texty.org.ua/d/fb-trolls/index_eng.html.

27 Zhdanova et al. pp. 16.

28 "Ukraine to Block Russian Social Networks." *BBC News,* May 16, 2017, sec. Europe. http://www.bbc.com/news/world-europe-39934666.

29 "Freedom of the Press 2017: Turkey Country Report." Freedom House, April 26, 2017. https://freedomhouse.org/report/freedom-press/2017/turkey.

30 Shearlaw, Maeve. "Turkish Journalists Face Abuse and Threats Online as Trolls Step up Attacks." *The Guardian,* November 1, 2016, sec. World news. http://www.theguardian.com/world/2016/nov/01/turkish-journalists-face-abuse-threats-online-trolls-attacks.

31 Kizilkaya, Emre. "Turkey's Ruling AKP Fields New 'digital Army'." Hürriyet Daily News, May 14, 2015. http://www.hurriyetdailynews.com/turkeys-ruling-akp-fields-new-digital-army.aspx?PageID=238&NID=82384&NewsCatID=550.

32 Arsu, Sebnem, and Dan Bilefsky. "In Turkey, Twitter Roars After Effort to Block It." *The New York Times,* March 21, 2014, sec. Europe. https://www.nytimes.com/2014/03/22/world/europe/turks-seek-to-challenge-twitter-ban.html.

33 "Twitter Transparency Report: Removal Requests." Twitter, June 2017. https://transparency.twitter.com/en/removal-requests.html.

[34] Morales, Silvia. "Feature: Turkey Trolls' Use of Insults Stifling Reporting." International Press Institute. Accessed October 9, 2017. https://ipi.media/feature-turkey-trolls-use-of-insults-stifling-reporting/.

[35] Shearlaw, Maeve. "Turkish Journalists Face Abuse and Threats Online as Trolls Step up Attacks." *The Guardian,* November 1, 2016, sec. World news. http://www.theguardian.com/world/2016/nov/01/turkish-journalists-face-abuse-threats-online-trolls-attacks.

[36] Cox, Joseph. "I Bought a Russian Bot Army for Under $100." *The Daily Beast,* September 13, 2017, sec. tech. https://www.thedailybeast.com/i-bought-a-russian-bot-army-for-under-dollar100.

[37] "Freedom of the Press 2017: Philippines," April 27, 2017.

[38] Ressa, Maria A. "Propaganda War: Weaponizing the Internet." Rappler, October 3, 2016. https://www.rappler.com/nation/148007-propaganda-war-weaponizing-internet.

[39] Hofileña, Chay F. "Fake Accounts, Manufactured Reality on Social Media." Rappler, October 16, 2016. https://www.rappler.com/newsbreak/investigative/148347-fake-accounts-manufactured-reality-social-media.

[40] Ressa, Maria. "How Facebook Algorithms Impact Democracy." Rappler, October 8, 2016. https://www.rappler.com/newsbreak/148536-facebook-algorithms-impact-democracy.

[41] Ressa, Maria. "How Facebook Algorithms Impact Democracy." Rappler, October 8, 2016. https://www.rappler.com/newsbreak/148536-facebook-algorithms-impact-democracy.

[42] "Poe Hits Mocha's Attempt to 'Reclassify' Rappler as Social Media." Rappler, November 9, 2017. https://www.rappler.com/nation/187881-grace-poe-mocha-uson-rappler-social-media.

[43] "Panaligan, Rey. "Rappler Faces Tax Evasion, Libel Charges at DOJ Probe." Manila Times, April 24, 2018. https://news.mb.com.ph/2018/04/24/rappler-faces-tax-evasion-libel-charges-at-doj-probe.

[44] "Freedom of the Press: Bahrain." Freedom House, March 10, 2016. https://freedomhouse.org/report/freedom-press/2016/bahrain.

[45] Perlroth, Nicole. "FinSpy Software Is Tracking Political Dissidents." *The New York Times,* August 30, 2012, sec. Technology. http://www.nytimes.com/2012/08/31/technology/finspy-software-is-tracking-political-dissidents.html.

[46] Marczak, Bill. "'Time for Some Internet Problems in Duraz': Bahraini ISPs Impose Internet Curfew in Protest Village." Bahrain Watch, August 3, 2016. https://bahrainwatch.org/blog/2016/08/03/bahrain-internet-curfew/.

[47] Anderson, Collin. "Dimming the internet: Detecting Throttling as a Mechanism of Censorship in Iran." June 18, 2013. http://arxiv.org/abs/1306.4361.

[48] Faris, Robert, John Kelly, Helmi Noman, and Dalia Othman. "Structure and Discourse: Mapping the Networked Public Sphere in the Arab Region," 2016. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2744915.

[49] Bolsever, Gillian. "Computational Propaganda in China: An Alternative Model of a Widespread Practice." Working Paper. Oxford, UK: OII, June 2017. http://comprop.oii.ox.ac.uk/2017/06/19/computational-propaganda-in-china-an-alternative-model-of-a-widespread-practice/.

[50] King, Gary, Jennifer Pan, and Margaret Roberts. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* 107, no. 2 (May) (2013): 1–18.

[51] Isaac, Mike. "Facebook Said to Create Censorship Tool to Get Back Into China." *The New York Times,* November 22, 2016, sec. Technology. https://www.nytimes.com/2016/11/22/technology/facebook-censorship-tool-china.html.

[52] Bolsever.

[53] Mozur, Paul, and John Markoff. "Is China Outsmarting America in A.I.?" *The New York Times,* May 27, 2017, sec. Technology. https://www.nytimes.com/2017/05/27/technology/china-us-ai-artificial-intelligence.html.

[54] Bolsever, Gillian. "Computational Propaganda in China: An Alternative Model of a Widespread Practice." Working Paper. Oxford, UK: OII, June 2017. http://comprop.oii.ox.ac.uk/2017/06/19/computational-propaganda-in-china-an-alternative-model-of-a-widespread-practice/.

[55] King, Gary, Jennifer Pan, and Margaret E. Roberts. "How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument." *American Political Science Review* 111, no. 3 (2017): 484–501.

[56] Botsman, Rachel. "Big Data Meets Big Brother as China Moves to Rate Its Citizens." *Wired,* October 21, 2017. http://www.wired.co.uk/article/chinese-government-social-credit-score-privacy-invasion.

[57] Chen, Yuyu, and David Y. Yang. "The Impact of Media Censorship: Evidence from a Field Experiment in China." Job Market Paper. Stanford University, November 2017. https://stanford.edu/~dyang1/pdfs/1984bravenewworld_draft.pdf.

[58] Clark et al.

[59] Xynou, Maria, and Arturo Filastò. "Internet Censorship in Iran: Network Measurement Findings from 2014-2017." OONI, September 28, 2017. https://ooni.torproject.org/post/iran-internet-censorship/.

[60] Larson, Christina. "China's Massive Investment in Artificial Intelligence Has an Insidious Downside." *Science,* February 7, 2018. http://www.sciencemag.org/news/2018/02/china-s-massive-investment-artificial-intelligence-has-insidious-downside.

[61] Williams, Greg. "Why China Will Win the Global Race for Complete AI Dominance." *Wired UK,* April 16, 2018. http://www.wired.co.uk/article/why-china-will-win-the-global-battle-for-ai-dominance.

Center for International
Media Assistance

NATIONAL ENDOWMENT FOR DEMOCRACY
1025 F STREET, N.W., 8TH FLOOR
WASHINGTON, DC 20004

PHONE: (202) 378-9700
EMAIL: CIMA@ned.org
URL: https://cima.ned.org

CIMA
CENTER FOR INTERNATIONAL MEDIA ASSISTANCE

NED
National Endowment
for Democracy
Supporting freedom around the world