Computational
Propaganda
Research Project

Online Supplement to Working Paper 2018.1 Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation

Samantha Bradshaw, *University of Oxford*
Philip N. Howard, *University of Oxford*

OXFORD
INTERNET
INSTITUTE

UNIVERSITY OF
OXFORD

Contents

METHODOLOGY NOTES

These are the background case notes complied for MEMO 2018.1: Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation. For details on the methods behind this content analysis please see the methodology section of the report. This document contains data from over 500 sources organized by country. The sources include high quality news articles, academic papers, white papers, and a range of other grey literature. As an annotated bibliography, the country cases here make use of significant passages from these secondary sources, and every effort has been made to preserve full citation details for future researchers. The full list of references can be found in our public Zotero folder, with each reference tagged with a country name.

ARGENTINA

Fake news frequently circulates on both social networks and the mainstream media in Argentina, and often includes political propaganda. For example, in 2017, a former member of Cristina Kirchner's cabinet tweeted a photograph of an injured woman, saying she was a teacher who had been hurt by the police while protesting in the Plaza de Congresos (National Congress Square), amid demonstrations related to the "Escuela Itinerante" (Itinerant School) conflict. The politician later apologized and deleted the tweet (Chequeado, 2017).

The website *Chequeado*—a digital non-profit fact-check agency—has a partnership with Facebook to address issues such as these and uses the hashtag #Falsoenlasredes ("Fakes on the Internet") to highlight fake news. This is part of a larger Argentinian initiative of "news literacy" which plans to give people the tools to be more informed about the news they are consuming (Jenner, 2018; Rost, 2017).

Fake profiles in social networks are also used by politicians and political parties. In 2012, a journalistic investigation conducted by the Argentine TV show *Periodismo Para Todos* ("News for everyone") revealed that 400 fake Twitter accounts were being used to post and disseminate messages in favor of Cristina Kirchner (The Observer, 2012). The election campaign that, in 2015, elected Mauricio Macri (the current president of Argentina), also made ostensive use of social networks, mainly with segmented advertising. It was also reported that Macri had 35 to 40 social media specialists working for his party to create trends and parrot government messages through fake accounts (Telesur, 2016). According to data from the Argentine Electoral Chamber, Macri spent 14 million pesos in segmented advertising in Facebook, Google and Yahoo! in the

presidency electoral campaign of 2015. In the same period, the Frente para la Victoria, an opposition party, spent 11 million pesos (Recalt, 2016).

A report produced by the consultancy Solo Comunicación has ranked Argentinian politicians by the number of fake followers on Twitter. The fake accounts were created to increase the number of followers and to share political propaganda in a disguised way. According to the report, the cost of the scheme was 1 dollar per 100 followers (Paladini, 2018). Many of these followers of political accounts are part of groups that seek to influence the process of public opinion formation. The accounts are used to increase the number of followers, praise or question government actions, show popularity, and generate interactions and false realities (Schurman & Romeo, 2018). The first reported case was in 2012, when the investigative journalist Jorge Lanata denounced Kirchner's aides as having created at least 400 Twitter accounts that were disguised as the accounts of regular citizens. According to Lanata, these bots simultaneously posted hundreds of messages in favor of Kirchner's decisions, and also launched aggressive attacks against journalists who disapproved of her government. The bots used profile pictures of actors and academics, had very few followers, and published dozens of tweets per day—only posting messages related to Cristina Kirchner and Argentine politics (Rueda, 2012). According to the engineer who led the investigation, Juan Carlos López, the bots used a special program to accelerate the creation of 'trending topics' in favor of Kirchner's government management (El Observador, 2012).

More recently, Solo Comunicación ranked the Argentinian leaders with the highest numbers of fake followers. Cristina Kirchner leads the ranking, with 2.8 million fake followers (53% of the total fake accounts following the ranked politicians). She is followed by Mauricio Macri, with almost 2 million fake users (41% of the total). Governors also show a high percentage of false followers: over 60% of the followers of the officials Alfredo Cornejo (Mendoza) and Gerardo Morales (Jujuy), and their opponents Omar Gutiérrez (Neuquén) and Juan Manuel Urtubey (Salta) are fake. Legislators such as Alfredo De Angeli and Federico Pinedo, as well as lower-ranking officials of the National Cabinet, such as Ricardo Giacobbe and Lucas Delfino, both have fake-follower rates exceeding 70% (Paladini, 2018).

Macri's 2015 campaign was based on significant investment in targeted advertising, especially on Facebook, and claimed to have been inspired by Obama's campaign in the United States. For example, if Macri planned to travel to the city of Río Cuarto, Córdoba, advertising on that day was aimed at male and female Facebook users (over the age of 18), living in Río Cuarto (Recalt, 2016). After the election, Telesur reported that President Mauricio Macri created a new

department in the presidential palace which employs people to control an "army of trolls," responsible for boosting his image and fire back at critics on social media (Telesur, 2016).

According to Freedom House's 'Freedom on the Net 2017' report, people were detained on suspicion of involvement in a hacking incident in 2016, in which the minister of security's personal Twitter account was used to send disinformation and insults. On other occasions, users were prosecuted and charged for issuing threats against the president and other public officials on social media (Freedom House, 2017).

The revelations regarding Cambridge Analytica and political manipulation also reached Argentina. In one of the videos released by the U.K.'s Channel 4, one of Analytica's directors briefly mentioned Argentina as one of the places in which the company had acted, but he did not go into details about dates and names of candidates and political parties. No further information about the allegation has been released so far.

AUSTRALIA

In Australia, leading politicians have significant followings of fake accounts on social media. Axel Bruns and Brenda Moon, researchers at the Queensland University of Technology, recognize social media as a fundamental part of political campaigning in present-day Australia. This development is situated in the wider context of popular adoption of social media—for instance, 95% of Australian Internet users use Facebook and 19% use Twitter. Moreover, the 'Reuters Institute Digital News Report 2016' found that 52% of Australians read online and social media news sources. Meanwhile, trust in mainstream news is low, which Bruns and Moon suggest is linked to the concentration of news ownership among only a few businesses, including Murdoch's News Corporation.

In May 2018, Labor Senator Kimberley Kitching's Twitter following was brought to the attention of the Australian Cyber Security Centre for being followed by thousands of Russian accounts. 7,000 (27%) of Kitching's followers were found to use Russian as their default language (according to ExportTweet). Separately, Sasha Talavera of Swansea University confirmed Kitching had 7,063 or 27.44% Russian followers. Whether Kitching or supporters had purchased these followers is unknown, but it shows that Russian accounts are present in Australian political campaigning on social media.

Kitching is not the first prominent politician to be in the news for having fake followers online. According to a study by Twitter Audit, nearly half (298,000) of the current Prime Minister

Malcolm Turnbull's followers are fake. Twitter Audit takes a sample of 5,000 followers and assesses ratios of real to fake followers based on the 5,000 accounts' follower to follows ratio, as well as the ratio of tweets posted to days since the account was created. The same study indicates Opposition Leader Bill Shorten has 36% fake accounts among his 155,000 followers and Foreign Minister Julie Bishop has 34% fake followers among her 195,000 followers. While the methodology has several limitations, it gives an impression of the presence of fake accounts in this space.

According to reports, Tony Abbott, prime minister of Australia from 2013 to 2015, had a significant fake Twitter following during his campaign for government. A 2013 Fairfax study found that 41% of his most recent 50,000 followers on Twitter were fake. On August 10, 2013, his following rocketed from 157,000 to 198,000. According to Australian reports, Abbott's following had increased by around 3,000 per day in the run-up to August 10. The Liberal Party conducted an internal investigation into the authenticity of his followers. A spokesperson commented that the surge was conducted by "an individual or party, not associated with the Liberal campaign." Reporters at *The Conversation* published a screenshot of Abbott bots in action. While denying involvement in the purchase of fake accounts for Abbott, the Liberal Party stated it was working with Twitter to remove the accounts. Julia Gillard, prime minister between 2010 and 2013, had the second largest proportion of fake followers with 40%. These metrics had been estimated by StatusPeople, which analyzes 50,000 followers. The *Australian Business Review* has challenged these findings and the validity of StatusPeople's methodology. In general, Twitter denounces external estimators as invalid estimators of fake accounts on the platform.

Turnbull's government has publicly presented evidence of Russian interference in the U.S. 2016 election as a watershed moment for Australia to advance its cyber defense infrastructure. In 2017, the Australian military launched its Information Warfare Unit, which is responsible for defending Australian military targets from cyber attacks and preparing for foreign attacks, such as psychological operations. Professor Greg Austin described this unit's establishment as one of the biggest shifts in Australian defense strategy. Currently, as reported, 100 staff work in the unit, which will grow to 900 staffers within 10 years. What is more, the Department of Defence has been licensed IBM's Watson for a 3-year period for roughly U.S.$3 million (AUS$4 million). Its specific purpose has not been publicly stated. However, this contract follows the Department's trial of the technology for a proof-of-concept psychological operations campaign. Reportedly, Watson will analyze intelligence and select targets.

AUSTRIA

Austria has a comparatively long history of digital media manipulation within internal party competition. As early as 2004, the ÖVP called on their campaigners to anonymously publish negative comments on websites. In 2011, the former chancellor Werner Faymann had to deal with a scandal about sponsored Facebook fans of his official fan homepage, though Faymann and his party the SPÖ always denied that they had purchased the accounts. Moreover, in 2014 it became public that the ÖVP Vienna had hired a media agency to author thousands of comments by sock puppets in 2009.

While the role of digital campaigns gained significant importance within the Austrian elections, digital media manipulation also continued to appear. In the 2015 Vienna elections, there were cases of astroturfing where political parties launched sites without clearly identifying themselves as the originators. In the 2016 presidential elections, Facebook—as well as several other websites—were used by the political camps of both main candidates, Alexander Van der Bellen (independent candidate) and Norbert Hofer (FPÖ), for negative campaigning and content promotion without identifying their party affiliation. In the 2017 federal elections, it became public that a political adviser paid by the SPÖ launched two heavily compromising and partly anti-Semitic Facebook pages about the chancellor candidate Sebastian Kurz from the ÖVP without identifying himself as the creator. This happened apparently without knowledge of the party leadership of the SPÖ. In the same election, a former ÖVP politician also anonymously used a Facebook page for dirty campaigning against the SPÖ chancellor Christian Kern. Beyond political campaigns, websites or newspapers and their online spin-offs, like unzensiert.at or Wochenblick, attained prominence. While presented as ordinary news portals or newspapers, these media have a close relationship to FPÖ without making this transparent. Moreover, they are accused of spreading a significant amount of junk news.

Looking at the role of microtargeting, all political parties used the microtargeting infrastructure provided by Facebook as well several other digital campaigning tools. The extent to which it was used was still smaller compared to, for example, the U.S., as the usage of complex voter profile databases is limited due to European and Austrian data protection laws.

AZERBAIJAN

Azerbaijan's oil wealth has allowed its government (commonly referred to as Baku, the capital in which the government is based) to fund large international projects and rebuild its military in recent years (BBC, 2018). This has also included extensive investment in social media manipulation (Geybullayeva and Grigoryeva, 2018). In particular, Azerbaijan's conflict with its

neighbor Armenia and breakaway territory of Nagorno-Karabakh has traditionally driven conflicts over information; for example, Baku has propagated online conspiracy theories regarding pogroms against Armenians (Kucera, 2018; News.az, 2011).

Reports of social media manipulation in Azerbaijan first emerged in 2011 and focused on the IRELI ("Forward") Youth (also called Ireli youth union or IRELI Youth Group) (News.az, 2011). IRELI Youth was formed in 2005, only a few months after a similar youth group in Russia (Safarova, 2018). It is affiliated with the government and was set up to "take active part in information war." They cultivated websites and blogs dedicated to contentious historical events such as "the Karabakh problem" (News.az, 2011). They have also been known to post abusive comments on social media; individuals are frequently targeted on Twitter and other social media platforms if they criticize the government. The use of bots has also been observed (Geybulla, 2016). Independent journalists and activists—such as the investigative journalist Khadija Ismayilova—are often the targets of intimidation campaigns based on illicitly obtained intimate images (Freedom House, 2018).

Volunteer work with IRELI is considered to be an entry point for roles in public administration (Geybulla, 2016). More recently, IRELI's profile has declined, in part due to controversies surrounding its leader's ties to the Fethullah Gülen movement, the controversial group which has been denounced by the current Turkish government (Safarova, 2018). Instead, the strategy seems to have evolved towards large Facebook groups which post predominantly positive messages about Azerbaijani history. It has been confirmed that these are funded by the government, although this is not publicly disclosed (Safarova, 2018).

BRAZIL

News media articles have reported the involvement of politicians and political parties in activities to manipulate public opinion over social media, either with the help of their own campaigning personnel or through companies hired to run social media campaigns, especially in the context of high-profile elections.

Secondary literature indicates that the operation and production of automated content has not occurred exclusively in one political pole. A study carried out by the Getulio Vargas Foundation (FGV) suggests that groups with different interests have been using automatically generated content to influence discussions on Twitter with the objective of generating an advantage for political actors. The research also identified bots operating abroad, suggesting the existence of actors beyond the national borders who were operating these mechanisms (Ruediger, 2017).

Another recent study looked at the role of online manipulation tactics during three political events: the presidential election in 2014, local elections in Rio de Janeiro in 2016, and the impeachment of former president Rousseff in the same year (Arnaudo, 2017). Various articles also showed that, during the 2014 presidential election, bots were used to promote both run-off candidates Dilma Rousseff (from the Workers Party – PT) and Aécio Neves (from the Brazilian Social Democracy Party – PSDB), particularly during TV debates between them.

According to an article published by the news agency A Pública, Xico Graziano, head of the PSDB's social media, admitted that the party campaign used fake accounts and robots in the 2014 campaign (Viana, 2015). The article also reported the existence of an internal memo written by the President's Office of Communication (Secom) leaked in March 2015, which mentions the use of bots used by the PT in the 2014 campaign. The document also stated that around 50 bots were used by the candidate from the opposition party (Aécio Neves, from PSDB) and that they continued to be used after the election, spreading messages criticizing the elected government—which helped to support the pro-impeachment campaign.

Computational propaganda has also been reported in local elections. According to Arnaudo, "in the elections for mayor of Rio de Janeiro in 2016, botnets appeared to be particularly active in the campaign. Marcelo Crivella, the right-wing leader of an evangelical mega-church, and Marcelo Freixo, a state representative, professor and member of the left-wing Socialism and Liberty Party, faced off in the final round. Both candidates accused each other of spreading online rumors and complained to elections authorities and in public debates about rampant fake news" (Arnaudo, 2017).

A BBC News investigation uncovered that, during Brazil's 2014 general election campaign, at least 100 fake accounts on Twitter and Facebook were active. According to the news article, a Rio-based PR company whose clients include a number of leading politicians was hired to run fake social media accounts to influence public opinion. A former worker reported having received around 360 U.S. dollars per month to run 20 fake accounts on Facebook and Twitter. His story is backed up by the accounts of three other young people who also worked as social media "activators" during the 2014 campaign. They all told BBC Brasil that they worked from home via Skype and often used the social media management platform Hootsuite to control many accounts simultaneously (Gragnani, 2017).

During the 2014 presidential election, the PT documented various accounts that appeared to be automated on Twitter, Facebook and other social networks attacking Dilma and supporting

Neves. The cases were linked to a businessman who received R$130,000 (Brazilian reais) to support the campaign (Arnaudo, 2017).

The leaked Secom memo estimates that PSDB spent around 10 million reais between November 2014 and March 2015 to manage the bots, create content supporting impeachment of the elected president, and sending WhatsApp messages to disseminate the content (Viana, 2015).

Brazilian political marketing companies have invested in the development of computer propaganda techniques and strategies. The Brazilian company *Face Comunicação On Line Ltda* received 130,000 reais from the PSDB between August and September 2014 to monitor social media and "sentiment analysis" of what was published on Twitter and Facebook during the electoral debate (Aragão, 2014).

Another company called Stilingue has developed software to monitor social media platforms, which is able to identify voters who are for and against a particular issue. The company, which has 50 employees in Ouro Preto and 10 in São Paulo, claims that two of its clients are pre-candidates for government positions in the 2018 elections (Mota, 2017).

In March 2017, the company A Ponte Estratégia announced a partnership with Cambridge Analytica, with the objective of adapting its profiling strategies to the Brazilian context. However, after the recent scandal involving Cambridge Analytica, which included allegations of illegal activities in Brazil, the CEO of *Ponte* announced the end of the partnership.

A recent investigation revealed that the federal deputy Fernando Francischini used a quota of R$24,000 (around U.S.$7,000) of his parliamentary allowance to make monthly payments to a media company known for running fake-news websites, between December 2017 and March 2018. The payments were made for the company Novo Brasil Empreendimentos Digitais and are recorded in six tax invoices made available on the transparency portal of the Chamber of Deputies. The websites managed by the company have published many stories praising the deputy Francischini and criticizing his political opponents (Toledo, 2018).

According to the BBC, the Rio-based PR company created fake accounts to create a buzz around its clients. Their strategy included praising whichever political candidates they were being paid to support, attacking their opponents and sometimes joining forces with other fake accounts to create trending topics. They also aimed at winning debates through sheer volume, by posting much more than the general public could counter-argue (Gragnani, 2017).

The FGV study, based on analysis of six case studies, also indicates the use of bots in the Brazilian political debate. According to the report, bots have been used by many different political parties not only to obtain followers, but also to conduct attacks against the opposition and to forge artificial discussions. "They manipulate debates, create and disseminate fake news and influence the public opinion by posting and replicating messages in a large scale" (Ruediger, 2017). According to the FGV research, bots were responsible for more than 10% of the interactions on Twitter during the presidential elections of 2014. Almost 20% of the interactions in the debate between users in favor of Aécio Neves during the second round of the 2014 elections was motivated by bots. During demonstrations related to Dilma's Rousseff's impeachment process in 2016, bot-provoked interactions represented more than 20% of the debate between supporters of the then president. Also, in the Brazilian general strike of 2017, more than 20% of Twitter interactions between users in favor of the strike were provoked by bots (Ruediger, 2017).

Arnaudo identified that modern campaigns linked together various social networks in a coherent strategy "using WhatsApp groups to drive people to more public forums on places like Facebook and Twitter." (Arnaudo, 2017). In Rio's local election, "researchers at the Federal University of Espírito Santo found a botnet of 3,500 accounts on Twitter attacking one of the candidates with repeated messages with the same phrase, often posting 100 or more times per hour." (Arnaudo, 2017).

A member of the Aécio Neves' digital team revealed to A Pública that one of the main techniques employed by his campaign in 2014 was "renting" famous profiles on Twitter—some with more than 1 million followers—by paying for their owners to post positive comments about the candidate. On Facebook, according to this former employee, the fake profiles were meant to interact with opponents, to intimidate critics (Viana, 2015).

A former member of Dilma Rousseff's campaign said she worked up to ten hours a day in partnership with a designer to produce posts that highlighted positive news about the candidate, or negative news about the opponents. In order to mask the source of the posts, the campaign employees used VPN connection, to make it difficult to identify the IP address used to post the content (Viana, 2015).

According to an investigation conducted by the newspaper *O Globo*, the right-wing political group Movimento Brasil Livre (MBL) and the website Ceticismo Político were responsible for boosting a campaign of fake news against Marielle Franco, councillor of a leftist party who was murdered in March 2018 in Rio de Janeiro (Cariello & Grillo, 2018).

Fábio Malini, coordinator of Labic (Laboratory of Image and Cyberculture) of the Federal University of Espírito Santo (UFES) states that 3,500 robots have made uninterrupted attacks on the candidate Marcelo Freixo on Twitter. The campaign of opposition candidate Marcelo Crivella denied any connection to the case (Albuquerque, 2016).

The Brazilian general election of October 2018 will be contested online as never before in the history of the country, and the use of the Internet will be crucial to the success of candidates and political parties. As this case study has showed, the use of the Internet by political campaigns has already happened, but it is likely to become increasingly more relevant. In 2018, for the first time, online paid advertisements will be permitted. Candidates will also be allowed to pay to boost posts, but it will be prohibited to use bots or fake profiles to increase the visibility of content.

In light of that, different Brazilian government bodies are seeking ways to address, monitor and punish the deliberate dissemination of fake news, and the use of other mechanisms that might unduly influence how citizens receive information, especially in the electoral context. Between 2014 and 2018, the Electoral Justice approved a series of new rules regarding electoral campaigns on the Internet. Moreover, more than 16 bills about fake news are being analyzed by the National Congress.

CAMBODIA

Prime Minister Hun Sen's Facebook page had almost nine million followers in 2017 and was ranked by global public relations firm Burson-Marsteller as the eighth most popular of any world leader (Paviour, 2017). He has since been accused of artificially boosting his popularity, by hiring foreigners to create fake accounts and increase the number of fans of his page (BBC News, 2016). A report published by a local news outlet showed evidence that Hun Sen might have arranged the purchase of fake likes (Nass & Turton, 2016).

Like in other Southeast Asia countries, prominent users of Twitter in Cambodia have recently been followed by large numbers of anonymous new account holders, fueling unconfirmed theories that political or commercial actors are exploiting the social media site (Reed, 2018). The government and the opposition often accuse each other of spreading false information, including recently leaked conversations about corrupt business dealings and politicians' infidelity (Hutt, n.d.). Prime Minister Hun Sen regularly accuses critical media outlets of spreading "fake news" (Lema & Wongcha-um, 2018). The term "fake news," however, is

routinely manipulated by politicians in order to stifle criticism against them (Dara & Baliga, 2018).

There is evidence of bots and fake accounts being used by politicians to boost their popularity on social media. In March 2016, local outlet *The Phnom Penh Post* reported that the majority of Hun Sen's recent likes came from foreign accounts. The single biggest groups of likes came from Indian accounts (255,692), with significant numbers also from the Philippines (98,256), Myanmar (46,368), Indonesia (46,368) and several others, according to the Post (Nass & Turton, 2016). It is suspected that these followers are not real. According to the report, companies running offshore 'click farms' might be behind them, in which low-paid workers create fake accounts to help bolster likes, followers and views on their clients' social media profiles (Nass & Turton, 2016).

In Cambodia, as reported in many other Southeast Asia countries, many journalists, academics and political figures appeared to have gained hundreds of new Twitter followers at the beginning of 2018. Prominent Twitter users in Thailand, Vietnam, Myanmar, Taiwan, Hong Kong and Sri Lanka noticed the same phenomenon—a surge in followers from anonymous, recently created accounts, adopting local sounding names but barely engaging on the platform (The Straits Times, 2018). In each country, the accounts use regionally authentic names, languages and profile photos to follow local influencers (Ruiz & Saksornchai, 2018).

The accounts were created in March 2018 and have since followed hundreds of Twitter users, but most have not tweeted or accrued any followers themselves (O'Byrne, 2018). Hundreds of the accounts with Cambodian names have followed a variety of Cambodia-based Twitter users, including the Ministry of Education, the Quick Press Reaction Unit, Australian Ambassador to Cambodia Angela Corcoran, the CNRP's Deputy Director-general of Public Affairs Kem Monovithya, and dozens of reporters at Cambodia-based news outlets (O'Byrne, 2018). For example, Maya Gilliss-Chapman, a Cambodian tech entrepreneur, said her Twitter account @MayaGC was being swamped by a daily deluge of follows from new users. She says she acquired over 1,000 new followers since the beginning of March (The Straits Times, 2018). Danielle Keeton-Olsen, an American journalist in Cambodia, said her Twitter followers surged from about 700 to over 1,700 in April 2018 (Ruiz & Saksornchai, 2018). Some affected users have speculated that one or more state actors might be behind the new accounts (Reed, 2018).

The proliferation of false information is a problem in Cambodia, as it is for other countries in the region (Salaverria, 2017). Recently, a government officer working for the Presidential Communications Operations Office (PCOO) was called out for sharing a post that misquoted

Canadian Prime Minister Justin Trudeau as saying that it was "not possible" for Canada to take back garbage that had earlier been shipped to the Philippines. In fact, Trudeau had said it was "now theoretically possible" for Canada to take back the trash (Salaverria, 2017).

Regarding surveillance and content removal, Cambodian citizens have frequently been in the sights of authorities over political speech on Facebook critical of the ruling party (Dara & Baliga, 2018). The government has recently created a task force to monitor the spread of "fake news" on social media platforms and through private text messaging. In a ministerial order signed on May 28, 2018, three ministries agreed to work with telecoms firms "to prevent the spread of information that can cause social chaos and threaten national security." (Vicheika, 2018). The Cambodian government is also considering new legislation targeting 'fake news', which has stoked fears of censorship, as many believe the new law could target those critical of the government. Local activists fear that, with the most critical media expelled from the country, the government may use an anti-fake-news law to target and manipulate civil society (Hoekstra, 2018).

According to Freedom House, politically motivated blocking has not yet been systematically applied in Cambodia, although it has been observed on a case-by-case basis. "Blogs blocked for supporting the political opposition, such as KI Media and Khmerization, were available through at least some ISPs during the coverage period, indicating that censorship orders are unevenly executed." (Freedom House, 2017).

Freedom House also reported allegations of paid content manipulation made in late 2016 involving an online activist and social media celebrity, Thy Sovantha, who claimed to have been offered U.S.$1 from representatives of the prime minister to lead campaigns against acting CNRP President Kem Sokha (Freedom House, 2017).

CHINA

The Chinese government often uses the term "fake news" to delegitimize criticism of the communist country. The discourse of fake news is used to crack down on dissident voices or to discredit opinions that confront the government. China's public officials want to "maintain public opinion control in order to maintain social and political stability." (Freedom House, 2018). According to *The Wall Street Journal*, since 2014, "while it didn't explicitly spell out what it meant by "fake news," the government has been cracking down on the dissemination of rumors or thinly sourced reports that it says contribute to social instability" (WSJ, 2014). According to the People's Daily, nine government departments will be involved in the crackdown on such activity (WSJ, 2014).

Regarding external activity, there have been reports of attempts to influence the recent U.S. presidential election from within China. According to WIRED, in 2016, Chinese-American blogger Xie Bin and seven others launched a WeChat page aimed at influencing Chinese-Americans to vote for Trump: "They called it "The Chinese Voice of America" (CVA) and published several articles each week that drew from right-wing websites in English, as well as concerns people shared in Mandarin in WeChat groups. Within months, it had more than 32,000 followers on WeChat. Even more people shared its content in private WeChat groups and commented about it on Chinese-language websites that center around WeChat content." (Gud, 2017).

Private companies that provide social media manipulation might also be operating in China. Since May 2017, more than 200 people in China have been arrested, and thousands of others confronted by police. Social media accounts and "illegal" websites have been seized as part of a campaign against organizations called *wǎngluò shuǐjūn* or "Network Navy" (網絡水軍—literally, "network water army"), that offer services that include boosting clients' websites on search engines for specific keywords along with general brand promotion and marketing (Gallagher, 2018).

There are reports that an obscure American company called Devumi sells Twitter followers and retweets to celebrities, businesses and anyone who wants to appear more popular or exert influence online. There is evidence that Devumi has more than 200,000 customers worldwide. According to *The New York Times*, an editor at China's state-run news agency, Xinhua, paid Devumi for hundreds of thousands of followers and retweets on Twitter. Even though the Chinese government has blocked Twitter in the country, it is widely used for propaganda abroad (Confessore, Dance, Harris, & Hansen, 2018).

Blocking of critical content has been widely employed by the government for many years, according to several sources. Internet service providers are required to block websites and delete content as instructed by censors. Thousands of websites have been blocked, many for years, including major news and social media hubs like *The New York Times*, *Le Monde*, Flickr, YouTube, Twitter, Instagram, and Facebook (Freedom House, 2018). For example, in mid-2017, two experimental "chatbots" from Tencent QQ, a popular Chinese messenger app (Baby Q and Little Bing) were removed after apparently voicing criticism of the government (Allen, 2017). More recently, in February 2018, the Communist Party proposed removing a clause in the Constitution which limits presidencies to two five-year terms—which means President Xi Jinping could remain as leader after the end of his second term in 2023. Because of this, several key terms

were subjected to heavy censorship on Sina Weibo, China's Twitter-like microblog. One of the censored phrases was "Winnie the Pooh" to avoid users making derogatory posts against President Xi, as it is a well-known nickname that social media users have coined for him (Allen, 2018).

China has a problem with fake news, too. In an interview with Bloomberg Television in 2017, Baidu Inc. president Zhang Yaqin said that China faces challenges similar to the ones faced by Western countries, despite operating one of the world's largest and most sophisticated online surveillance machines (Bloomberg News, 2017). In recent years, media outlets have reported an increase in fake news on social media platforms, especially in more private communities on WeChat, Telegram, Signal, and WhatsApp, where information is harder to track and verify. In general, WeChat's design does not make it easy to fight bias or fake news. Information on the platform spreads quickly within and between WeChat groups, but the sources of information—and therefore their verifiability—are de-emphasized, to the extent that sources are almost completely ignored (Gud, 2017).

As Freedom House has reported, online content is subject to extensive manipulation, with propaganda officials instructing Internet outlets to amplify content from state media (Freedom House, 2017). In 2009, news outlets reported that the Chinese Communist Party had raised a "50-Cent Army" of astroturfers who were paid RMB0.50 for each patriotic pro-Chinese comment they post on blogs and social media sites. Some estimates have the size of the army at 300,000 people (Doctorow, 2009).

In 2014, leaked documents detailed how the Chinese government employed people to post pro-government messages on the Internet, as part of a broader effort to "guide public opinion." Among the leaked documents were instructions to paid commenters, their posting quotas, and summaries of their activity. The emails reveal hundreds of thousands of messages sent to Chinese microblogging and social media services like Sina Weibo, Tencent, and various Internet forums, including working links to the actual posts (Sonnad, 2014). According to a research published by King et al. in 2017, rather than debating critics directly, the Chinese government tries to derail conversation on social media it views as dangerous.

In a large-scale empirical analysis of this operation, King et al. estimate that the Chinese government fabricates and posts about 448 million social media comments a year. The researchers claim that the Chinese regime's strategy is to avoid arguing with skeptics of the party and the government, but rather that the goal of this massive and secretive operation is to distract the public and change the subject, as most of these posts involve cheerleading for China,

the revolutionary history of the Communist Party, or other symbols of the regime (King, Pan & Roberts, 2017). The government does not refute critics or defend policies; instead, it overwhelms the population with positive news in order to eclipse bad news and divert attention away from actual problems (Illing, 2017).

More recently, since May 2017, there has been evidence of media accounts and "illegal" websites managed by the previously mentioned "Network Navy" (Gallagher, 2018). Network navies are loose organizations of hundreds or thousands of people recruited through sites targeted at "leisure workers," similar to Mechanical Turk jobs. They generally offer services that include boosting clients' websites on search engines for specific keywords along with general brand promotion and marketing. But they can also generate "press releases" and set up channels for getting fake news releases onto major Chinese mainstream media sites—sites designated by the Chinese government as approved news sources. They also offer followers to amplify messages on social media services such as WeChat, the Weibo microblogging site, Dianping (like Yelp), and RenRen (similar to Facebook) (Gallagher, 2018).

These network navies are also reportedly involved in the creation of spam email campaigns, fraudulent news sites, and social media trolling campaigns to shape public opinion. Another profit center for network navies is the deletion of negative posts on social media sites by aggressive use of sites' moderation flagging, by hacking, or paying off insiders with administrative access to various platforms to delete the posts. Usually these services target consumer complaints against a particular company. Network navy salespeople usually double the price for "content-sensitive posts," making them highly profitable (Gallagher, 2018).

Over the past two years, the Chinese government has set up initiatives to encourage the "good netizen" who spreads positive messages about China, such as the Communist Party youth league's "Volunteer Campaign to Civilise the Internet" (Yang, 2017). This has encouraged the organization of a nationalist volunteer social media army, known as 'little pink,' or *xiao fenhong*, a name derived from the color of a popular online forum used by nationalists (Yang, 2017). These young nationalist volunteers usually spread positive messages about China, often focusing on pop culture to whip up support, but also to coordinate 'mass bombings' of public figures' social media platforms, flooding targets with intimidating posts and shutting down online debate (The Economist, 2016; Yang, 2017). Their targets are varied, from Taiwan's pro-independence president to international airlines accused of mistreating Chinese customers. Lady Gaga's Instagram account was targeted last year after she met the Dalai Lama, the exiled Tibetan spiritual leader whom Beijing denounces as a separatist. Attacks, though usually spontaneous, are meticulously organized in reaction to perceived slights against China. The trolls share tips on

how to access Facebook, Twitter and other foreign sites blocked by Chinese censors (Yang, 2017).

As widely reported by several sources, the Chinese government has long been suspected of hiring as many as 2,000,000 people to insert huge numbers of pseudonymous and other deceptive comments into social media posts, as if they were the genuine opinions of ordinary people; the '50-Cent Army' already mentioned (King et al., 2017). King et al. identified 43,000 government-sponsored posts and comment campaigns, and estimated that about 1 out of every 178 social media posts is sponsored by the government, amounting to some 448 million posts per year (King et al., 2017).

More recently, network navies have been offering services online. Getting a post deleted costs from 300 to 3,000 yuan (about U.S.$50–$500). According to Ars Technica, one website operator said he made about 4,000 yuan (U.S.$636) a month deleting comments, which were mostly consumer complaints about product quality (Gallagher, 2018). In July 2017, the Chinese police arrested 77 suspected members of the network navy and seized nearly 4 million yuan (about U.S.$640,000) as well as computers, mobile phones, flash drives, and bank cards. Since then, there have been more than 40 coordinated operations by Chinese police agencies and over 100 million yuan (about $16 million) in cash seized. A CCTV reporter found over 2,300 network navy "shops" online, selling "news service" access (Gallagher, 2018).

With regard to the "little pink" army, many of them belong to the "Emperor's Board," an online forum followed by 29 million people, where "crusades" are coordinated. China's troll army also organizes via private groups on Facebook. The most popular of these has 40,000 members, who must express their support for the party (Yang, 2017). According to the *Financial Times*, the Communist Party provides support for the little pinks, arming them with memes produced by state agencies as well as private studios (Yang, 2017).

Another actor operating in China is the American-based company Devumi. Most of the Twitter accounts managed by Devumi resemble real people, and some are even associated with a kind of large-scale social identity theft. At least 55,000 of the accounts use the names, profile pictures, hometowns and other personal details of real Twitter users, including minors, according to *The New York Times* (Confessore et al., 2018).

Regarding fake news, in 2016, WeChat allegedly disabled more than 1.2 million links, deleted over 200,000 articles of alleged false information, and fined 100,000 accounts that either created or spread rumors, according to the Cyber Administration of China (Gud, 2017).

COLOMBIA

Several media outlets reported the spread of disinformation and rumors in online platforms related to the negotiations between the guerrilla group FARC and the Colombian government in the lead-up to a national referendum on a peace deal in October 2016 (Freedom House, 2017). After being rejected in the referendum, the accord was successfully renegotiated in November 2016, but it remains a highly divisive issue in Colombia and has led to increased levels of intolerance and polarization in public debate (Misión de Observación Electoral, 2018).

The political polarization surrounding the accord has also affected the media coverage of the 2018 electoral campaign, leading to a barrage of fake news and disinformation attacks against presidential candidates. According to BBC Monitoring, many of these attacks have sought to harness widespread antipathy felt by many in Colombia against the FARC, especially due to the lenient treatment received by former guerrillas in the agreement (BBC Monitoring, 2018).

After it converted into a political party, the FARC initially had a candidate for the presidential race, but later withdrew the candidacy. According to BBC Monitoring, "Colombian media outlets like weekly magazine *Semana* and independent website *La Silla Vacia* reported that social media trolls and disinformation peddlers were using the accusation of association or affinity with the left-wing guerrilla group to try to tarnish and discredit their targets in the eyes of voters." (BBC Monitoring, 2018).

In an attempt to curb the dissemination of fake news, the Constitutional Court of Colombia ruled in a judgement issued in January 2018 that any citizens, journalist or media outlet, must back up information with valid sources and not generate false news that could harm the reputation of a person—information should be grounded in truth and not just rumors. According to the court, although freedom of expression is a fundamental right, the rights of others should not be violated and the creation of fake accounts that can harm people might be considered a crime in Colombia (Colprensa - La República, 2018).

The Colombian organization Misión de Observación Electoral - MOE (MOE, Electoral Observation Mission) published a report about social media behavior during the 2018 campaign to Congress, which showed a high degree of intolerance in political speech. MOE reports that one in five messages posted in Facebook, Twitter, Instagram, YouTube and blogs related to the elections were charged with intolerance, aggressiveness and polarization—drawing on a data set of 8,188,417 messages analyzed (Misión de Observación Electoral, 2018a).

Also, according to MOE's report, the main arguments that were used to generate attacks in social networks were related to the public rage against the FARC, with 34% of the conversation relating to intolerance, followed by corruption with 26%, "Castrochavismo" with 24%, and paramilitarism with 7%. Related to that, most of the mentions of intolerance in social networks were attacks on FARCs candidates to the Congress, such as Rodrigo Londoño (Misión de Observación Electoral, 2018a).

Andrés Sepúlveda, a Colombian hacker, said in an interview with Bloomberg in 2016 that his first hacking job was breaking into the website of a rival of President Uribe in 2005, to steal a database of email addresses, and to spam the accounts with disinformation. He was paid U.S.$15,000 in cash for a month's work (Robertson, Riley, & Willis, 2016).

During the 2018 presidency campaign, the weekly magazine *Semana* reported that a Facebook group called 'People United against Communism and Socialism' had been publishing posts attacking presidential candidates Gustavo Petro, German Vargas Lleras, Sergio Fajardo, and Humberto De La Calle, by describing them as "communists and socialists" (BBC Monitoring, 2018). Gustavo Petro, the leftist candidate, alleged he had been slandered by claims he would expropriate Colombian businesses if elected, and used Twitter to ask for the attorney-general to investigate the authorship of this "fake news" (Rathbone, 2018).

Colombiacheck (a Colombian fact-checking agency) also reported that supporters of right-wing election frontrunner Ivan Duque had created websites, as well as Facebook pages and groups, to publish false information about Petro and Vargas Lleras. Similarly, they report that Petro supporters had set up their own pages to disseminate false and biased information against their candidate's election rivals (BBC Monitoring, 2018).

Social media sites have been used to spread misinformation about the referendum on the peace deal in Colombia. One example of a fake story in that context was the rumor that, according to a new law ('Roy Barreras'), pensioners would need to pay over 7% of their pensions to support the demobilized guerrillas ('"Roy Barreras", la ley que se "aprobó" en las redes sociales", 2016).

During the 2018 Congress campaign, social media spread accusations by the 'People United against Communism and Socialism' group accusing candidate Fajardo, from the political party Alianza Verde (Green Alliance), of being a "guerrilla member." This was later refuted by the Pinochometro, a local fact-checking tool that evaluates and checks information shared on social media. Similarly, the news outlet *La Silla Vacia* has recently analyzed a social media thread which

had alleged that Fajardo's Alianza Verde was the "FARC's 'plan B' to impose '21st century socialism' in Colombia in alliance with Gustavo Petro" (BBC Monitoring, 2018).

In a recent case investigated by Colombiacheck a Facebook posting accused presidential candidate De La Calle, who worked as the government's lead negotiator in the peace talks with the FARC, of being a "narcoterrorist" and "friends" with former FARC leaders (BBC Monitoring, 2018).

Regarding content moderation, there have been claims that Facebook has shut down a number of pages and member profiles belonging to the FARC for allegedly violating Facebook's community standards, even though the organization is no longer classified as a terrorist group ('Facebook Keeps Fake News but Censors Colombia's FARC', 2016).

CUBA

Media outlets have reported events in which the government-controlled Cuban mass media has generated disinformation through a mix of censorship, propaganda and limited access to the Internet (Gámez, 2017).

Freedom House has reported that news reports are often manipulated and biased according to the political priorities of the state (Freedom House, 2017). For example, several reports in Cubadebate have mentioned an alleged plan by the Pentagon's Miami-based Southern Command to topple Maduro. A document titled 'U.S. Southcom Operation "Venezuela Freedom," American Strategy to Overthrow the Maduro Government' has been posted on several web pages (Gámez, 2017).

There is also evidence of interference from foreign governments in political debate. In 2014, news outlets reported that the U.S. government had developed and implemented an app aimed at undermining the Cuban government. According to the news articles, the U.S. Agency for International Development (USAid) launched the app ZunZuneo, a social network built on texts. "According to documents obtained by the Associated Press and multiple interviews with people involved in the project, the plan was to develop a bare-bones "Cuban Twitter," using cell-phone text messaging to evade Cuba's strict control of information and its stranglehold restrictions over the internet." (Associated Press, 2014).

Surveys show that Cubans do not usually use social media to organize large-scale campaigns around political objectives, but rather for social networking—and even then in a limited way, as

Facebook and other platforms are only accessible in some Wi-Fi hotspots (Freedom House, 2017). Dissidents have rather used independent blogs to criticize the government, sometimes using the platform offered by the government itself, Reflejos, where content can only be published from a Cuban IP address (Dirección de Comunicación Institucional Joven Club, 2015). Human rights activists have reported the use of technical tools to manipulate public debate. The Foundation for Human Rights in Cuba (FHRC) has denounced the growing use of digital tools of cyber warfare against political dissidents in Cuba. They reported situations in which their email and Facebook accounts were hacked, and have reported more than 14,000 viral attacks on their websites. The objective, according to FHRC, is to generate or exacerbate conflicts among various organizations and to discredit them:

> They are resorting to the techniques of modern black propaganda: They falsify statements, edit video and audio tapes, and make photo montages that are then disseminated from the computers, phones, sites, emails and Facebook hijacked accounts of the opposition activists that they want to discredit. In addition, sites dedicated to "black propaganda" and psychological warfare have multiplied. These blogs, often under the facade of fictitious names, provide a platform for state security agents charged with spreading rumors, attacking the credibility of those who they find "uncomfortable," and sowing disinformation lines that justify the repressive operations of their institution (FHRC, 2017).

In October 2017, several Facebook accounts of opponents of Castro's government were hacked, according to Carlos Amel Oliva, youth leader of the Patriotic Union of Cuba, UNPACU (Martí, 2017).

There is also evidence of the use of bots and trolls by the Cuban government. Experts and activists have tracked dozens of automated social media accounts attempting to masquerade as humans, which are used to amplify certain hashtags and messages to influence what is trending. One strategy employed by them is the use of pictures of white attractive public figures (Torres & Vela, 2018). Students loyal to the Communist Party are allegedly being used as social media marketers, to amplify messages supporting the government. A local news outlet has reported that students from the University of Information Science in Havana are responsible for spreading socialist propaganda on Twitter, during events they referred to as the "Twitazo." Cubadebate has also evolved into an international effort to spread political propaganda in seven languages (Torres & Vela, 2018).

Regarding the app ZunZuneo, documents show the U.S. government planned to build a subscriber base through "non-controversial content" and then introduce political content aimed at inspiring Cubans to organize demonstrations against the regime. According to Associated Press, "at its peak, the project drew in more than 40,000 Cubans to share news and exchange opinions. But its subscribers were never aware it was created by the U.S. government, or that American contractors were gathering their private data in the hope that it might be used for political purposes." (Associated Press, 2014).

Content blocking and takedowns are also employed by the government to control political debate. According to Human Rights Watch:

> "the government controls virtually all media outlets in Cuba and restricts access to outside information. A small number of journalists and bloggers who are independent of government media manage to write articles for websites or blogs, or publish tweets. The government routinely blocks access within Cuba to these websites and only a fraction of Cubans can read independent websites and blogs because of the high cost of, and limited access to, the internet" (Human Rights Watch, 2018).

According to Freedom House's report, state-owned cell-phone provider Cubacel had been systematically filtering domestic SMS containing keywords such as "democracy," "dictatorship," and "human rights." (Freedom House, 2017).

The government has also attempted to exert control over the digital landscape by trying to direct popular demand for videos, games, and online social networking to government-controlled platforms. The Cuban government launched its own copycat versions of popular websites such as Wikipedia, Twitter, and Facebook, which can be more easily monitored and censored. According to Freedom House, "In 2010 the government launched Ecured, a copycat version of Wikipedia, and in 2013 they launched the social networking site La Tendedera, which is accessible from youth centers. In March 2015, the government launched the blogging platform Reflejos, where content can only be published from a Cuban IP address." (Freedom House, 2017).

ECUADOR

According to the organization Fundameios, the growing spread of misinformation comes from both government and opposition political actors. Both sides have been involved directly and indirectly in the design and dissemination of falsified, altered and decontextualized content, in order to confuse the population (Fundamedios, 2017).

The 2017 cyber troops report has already described the functioning of website Somos+, used by the government to investigate and respond to social media users who criticize Correa's administration. There have also been several reports of state-sponsored 'troll farms' in Ecuador. An investigation conducted by Fundación Mil Hojas in 2015 revealed that Correa had hired businesses to run "troll centers"—offices where social media users with fake accounts were paid to voice support for President Correa and attack his opponents (Alpert, 2018). Catalina Botero, former Special Rapporteur for Freedom of Expression for the Inter-American Commission on Human Rights, has reported that investigations had tracked the IP addresses of such troll farms and linked them to computers in government offices (Freedom House, 2017).

Fernando Balda, a deputy from the opposition party Sociedade Patriótica, has also accused Correa's government of setting up a troll center in order to harass journalists and critics of the government through false accounts on Twitter, Facebook and YouTube. Such accusations were corroborated by *Diario El Comercio*, a newspaper established over 100 years ago. The complaint involves high-ranking officials such as Fernando Alvarado, Secretary of Communication (Lara-Dillon, 2012). On the other hand, political opponents of former President Rafael Correa have also employed defamatory social media campaigns (Digital Guarimbas, 2017).

In November 2016, leaked documents related to publicist Kenneth Godwin revealed a proposal to use public funds to hire companies (Inteligencia Emocional and Kronopio) for the creation and expansion of political propaganda supporting the Ecuadorian government. These companies maintained a close relationship with Vinicio Alvarado Espinel, a high-ranking government official. According to the documents, Alvarado took U.S.$81,915.76 in commissions related to contracts made by Godwin with the Presidency (Ecuador Transparente, 2016).

Among the leaked documents were also a budget for operating propaganda accounts on social media and a proposal to handle social media campaigns attacking opposition leaders like ex-Secretary for Communications Mónica Chuji, local press watchdog Fundamedios, the Inter-American Commission on Human Rights and its Special Rapporteur for Freedom of Expression Catalina Botero, among others (Ecuador Transparente, 2016). Another document proposed the management of a 24/7 community manager and other security implementations for the monthly fee of U.S.$15,000, which included the creation of a troll center to influence public opinion on social media by, among other tactics, attacking and harassing dissidents (Ecuador Transparente, 2016).

On one occasion, the former president used a speech to name and shame people who had written abusive comments about him on Twitter and Facebook, revealing three people's full names, ages and addresses. He also asked his supporters to send thousands of tweets to each of them. In another situation, Correa attacked the satirical Facebook page Crudo Ecuador, which posted pictures and memes sending up the government (BBC News, 2015). The humorist behind the page was subsequently doxxed on social media (Viñas, 2018).

In 2017, in the lead-up to presidential elections, Freedom House reported that social media accounts belonging to politicians, journalists and opposition activists were hacked and used to disseminate messages against the opposition's vice-presidential candidate Andrés Paez (Freedom House, 2017; Puente, 2017).

Content takedowns were also used as a strategy to curb political discourse. A study by Fundamedios revealed that, between April and July 2016, approximately 30 Twitter accounts linked to anti-government users with high numbers of followers were suspended after receiving repeated complaints (Freedom House, 2017).

EGYPT

The Egyptian government maintains a repressive hold over information freedom, particularly on the Internet. These controls are situated amid an interplay of geopolitical tensions and a government which, following the overthrow of President Mohamed Morsi, pursues policies focused on political stability. President Abdel-Fattah el-Sissi's regime has violated human rights in the pursuit of this stability. Internet freedom is limited and reportedly over 400 prominent websites, many of them news websites, are blocked as of October 2017 (Freedom House).

Internet penetration in Egypt is around 39%. The government has centralized the Internet infrastructure and fiber-optic cables, creating highly controllable choke points (Freedom House, 2017). A notable example of the exercise of this power was the blocking of VoIP services on apps like WhatsApp, FaceTime, Skype and Facebook Messenger for one day, speculated to have been linked to the 3 month state of emergency that followed the terrorist attack at a church on Palm Sunday, April 9, 2017.

In preparation for the March 26–28, 2018 presidential election, the government warned against the spread of fake news on social media and of legal action against individuals who "undermin[e] that country's security," according to the Public Prosecutor Nabol Sadek. Sadek ordered his staff to monitor fake news on social media, which he called "forces of evil" that seek "to disturb the

public order and terrorise society." Al Monitor commented that Sadek's words repeated President Sisi's statement in January, where he called political opponents who called for an election boycott "evil forces." Al Monitor suggested Sadek's statement was directed at foreign media which have been critical of the Egyptian government, including the controversial BBC 'The Shadow over Egypt' report that aired on February 23, 2018. The government's offensive against fake news is situated in the wider context of blocking access to hundreds of news websites. In September 2017, according to the Cairo-based NGO, the Association for Freedom of Thought and Expression, over 420 websites have been banned. News banning has been a hallmark of the government's control of political voices in the run-up to the election. The Public Prosecutor launched a fake-news hotline on WhatsApp for citizens to report "news relying on lies and rumours" in both traditional media and on digital platforms. Fake news, in this case, was spoken of in terms of a threat to national security.

The government offensive on fake news has been leveraged against journalists. 18 journalists were arrested in 2017 and, by March 2018, at least six journalists had been arrested under the justification of fake-news containment. Harassment of journalists has complemented the shutdown of news websites as part of the wider strategy of online media censorship. The U.N. has denounced Egypt's crackdown on political freedom and press freedom before the presidential election. In May 2018, the activist Amal Fathy was detained for 15 days for inciting the overthrow of the ruling government, for "publishing lies and misusing social media," according to Reuters. Amal had posted a video on Facebook of her criticism of the prevalence of sexual harassment in Egypt, the government's failure to protect women, as well as the deterioration of human rights in Egypt. According to Amnesty International, the arrest was carried out in response to her damaging of the government's reputation. Amal's husband, Mohamed Lotfy, the Director of the Egyptian Commission for Rights and Freedoms, was arrested, too. Amnesty's analysis of Amal's video, however, concludes that the video does not contain incitement in any form. Amal was denounced in the mainstream media as an "April 6" activist.

Access to social media has been limited during periods of government shutdown of access to social media apps or removed content. For instance, a Facebook page under the name of President Sisi, which has over 800,000 followers, was closed down following its posting of a satirical image showing Sisi's mobile phone displaying 12 missed calls from King Salman of Saudi Arabia after the Egyptian court ruling not to cede two Red Sea islands (Freedom House, 2017). The government has said it has shut down hundreds of social media pages; in December 2016, a reported 163 Facebook pages were taken down and 14 administrators arrested for "inciting

people to commit acts of vandalism against state institutions and citizens" (Daily News Egypt; Freedom House, 2017).

Social media takedowns and arrests add to the chilling effects of the government's 2015 anti-terrorism law, which has criminalized the publication of information about militant attacks contradict government narratives. Authors face punishment of up to two years in prison. What is more, providing an alternate source of news is becoming increasingly difficult; for instance, in order to register a local domain, one must submit one's personal data and copies of one's national ID. The provision of such personal data discourages websites from criticizing politicians or the government. Additionally, news outlets that are exclusively digital are not recognized as official news outlets by the government, which makes obtaining access to people, sources or events difficult for online reporters (Freedom House, 2017).

In April 2017, Member of Parliament Rayed Abdel Sattar proposed a social media draft law in order to control Facebook pages, or users who spread fake news or terrorist content. The draft law's first article states that it will affect all social networks and apps connected to the Internet which are used as a means of communication; its second article states that the government will establish an authority that registers and permits both citizens and foreign residents to use social networks; and the third article states that once signed into law, citizens have six months to register. Citizens caught using social networks for communication without registration will be sent to trial and if found guilty face six months in prison or a fine of approximately U.S.$280 (LE5,000).

In addition to news blocking, the government has also used phishing techniques to target NGOs focused on human rights in Egypt. According to Citizen Lab, large-scale phishing campaigns have targeted at least seven human rights NGOs in what has been termed the 'Nile Phish', comprising at least two phases. The sponsor of the attacks, so far unknown, has demonstrated an interest in and familiarity with the activities of Egyptian NGOs. Citizen Lab observed a case where, within hours of the arrest of lawyer Azza Soliman in December 2016, phishing emails containing a false copy of her arrest warrant were sent to her colleagues.

GERMANY

While in many countries social media manipulation is a prevalent and dominating phenomenon in the political discourse, in Germany it plays a subordinate role. The only political actors that are regularly accused of social media manipulation and political disinformation are the German right-wing party AfD (Alternative for Germany)—the party responsible for a significant

proportion of junk news spread during the 2017 German federal elections. Moreover, there are reports about forms of astroturfing in terms of detailed coordination and cooperation between the party and an unofficial supporter network on Twitter. Other reports about cooperation with groups like Reconquista Germania—an organized, hierarchical group of up to 5,000 extreme-right activists who systematically pushed and supported AfD content on social media—only exist for minor party officials.

The role of microtargeting or dark advertising on social media is not a common phenomenon due to Germany's electoral and data protection laws. While the creation and use of complex voter profile databases is limited due to European and German data protection laws, in the 2017 elections all major German parties used the microtargeting infrastructure provided by Facebook. Nevertheless, the extent to which it was used was much smaller compared, for example, to the U.S., as social media in general plays only a subordinate role in news consumption in Germany.

Looking at the official state level, the creation of a new organization structure within the German Bundeswehr (the national army) called 'Cyber and Information Space' attracted a significant amount of attention. It was accompanied by several official strategy papers by the Federal Ministry of Defence and the Federal Ministry of the Interior which explicitly name information warfare as a central security challenge German security organs are faced with. Moreover, corresponding to counter measurements were part of several military exercises on the NATO and EU level, which the Bundeswehr was involved with. However, in general it appears the focus of the Cyber and Information Space structure within the Bundeswehr is more on classic cyber security issues like network security or counterintelligence and less on information warfare. There are no reports that the Bundeswehr is creating significant capacity in this area.

HUNGARY

Most efforts around computational propaganda in Hungary appear to be concentrated around the ruling Fidesz party and its leader Viktor Orbàn, the country's prime minister, who has held office since 2010. The general media environment seems to be heavily influenced by the government: according to Marius Dragomir, director at the Centre for Media, Data and Society at the Central European University's School of Public Policy, around 90% of Hungary's media is either directly or indirectly controlled by Orbàn's party, after the 2008 crisis exposed many media organizations open to takeovers.

Propaganda efforts online seem to be mostly designed in order to further political aims, and are both government-managed, government-funded and government-friendly. These include financing an entire fake-news industry, with a network of disinformation websites that both praise the current government and attack the opposition media as 'fake news', and coordinating campaigns against Hungarian-born billionaire philanthropist George Soros through billboards and Facebook advertising. Narratives emphasize antagonism along liberal vs. conservative lines, a conflict that mirrors Russian/Western antagonism in the region. Although there is little coverage about the matter, one of the debates currently in the press concerns precisely the extent of Russian influence, very much like in the other Visegrad group countries. Observers have noticed that in 2013–2014, disinformation came from Russian sources in terms of content and tactics, including the anti-Soros campaign and the strategic use of leaks against him which are reminiscent of the Russian DC Leaks site attributed to Russian cyber-espionage group Fancy Bear. From 2015 on, however, a more 'home-grown' industry has appeared that serves both propaganda and business purposes. A 2016 investigation disentangled Hungary's web of pro-Kremlin niche websites and found that many had lost readers and ceased publishing. In fact, it appears that pro-government disinformation matches Kremlin narratives without any direct influence from Russia. This is particularly flagrant in the context of the migration crisis, which fostered anti-EU stances closer to the positions of the Kremlin, often with fabricated stories (such as the fake story of the Swedish woman who had moved to Hungary because she no longer felt safe due to the presence of Muslim migrants).

INDIA

According to the 2017 Freedom House report on India, India enjoys an Internet penetration rate of 33%, a mobile penetration rate of 92%, and has emerged as the second largest Internet consumer base in the world with 431 million users (Freedom House, 2017). Of those, 21.6 million had fixed-line Internet connections. The last year has seen social media apps blocks, content blocks, and the arrest of content producers. Internet shutdowns and social media blocks have occurred frequently (37 reported shutdowns were ordered by local authorities), especially in the Jammu and Kashmir (J&K) regions, some of which have lasted several months. In Kashmir, in April 2017, 22 social media sites including Facebook, Twitter and WhatsApp were blocked. Shutdowns often followed violent protests over caste quarrels; for instance, in Haryana, shutdowns followed protests by the Jat caste over their eligibility for government affirmative action quotas (Freedom House, 2017).

The prime minister of India, Narendra Modi, has a highly visible footprint on social media; he is second to President Donald Trump as the most followed politician on Twitter. Reportedly, 60%

of his followers are fake according to Twitter Audit, which has a questionable assessment methodology. It is an external tool that was not built by Twitter; it takes a sample of 5,000 followers and makes an assessment based on the number of tweets, followers, mutual follows as well as other metrics if the accounts are fake or not. Twitter has denounced the methodology as "deeply flawed." Modi also has his own app, NMApp, which is a platform for communicating with his followers. Modi has been criticized for evading media scrutiny by communicating via the app.

Modi has been under fire for following Mahesh Vikram Hedge, the editor of *Postcard News*, a fake-news site that consistently praises Modi and the Bharatiya Janata Party. Hedge was arrested in March 2018 for publishing fake news that Muslims had attacked a Jain monk, who actually had been injured in a road accident. Hedge shared the news on Twitter, too. Modi's following of Hedge on Twitter suggests tacit support for fake news and intimidation in digital politics. Politicians in the Bharatiya Janata Party were criticized for using bots to amplify their messages on social media before they won the 2014 election. Freedom House reports that women, journalists and political activists face frequent trolling and violent threats; for instance, Gurmehar Kaur, a 20-year-old student in New Delhi, was sent rape and murder threats after she criticized hard-line right-wing students in a video in February (Freedom House, 2017).

Fake news that is spread on social media often aims at worsening existing religious and caste tensions. One video that has been in circulation on WhatsApp, for instance, is of a Hindu Marwari girl being beaten by Muslim men in Andhra Pradesh. The video also contains text explaining that the men attacked the woman for refusing to wear a burqa after marrying a Muslim man. The video image is shaky and it can be hard to discern details. However, this video was not filmed in India, but in fact shows a 16-year-old girl in Guatemala being attacked, reportedly for her suspected involvement in a murder. However, it is near impossible to know this without any mention of it and the accompanying fake text. It is disinformation that deliberately seeks to provoke anger and fear between religious groups. Its falsity was exposed by Sinha, a co-founder of Alt News, an anti-fake-news website that exposes and challenges fake news in India.

In March 2018, Smriti Irani, the Union Minister for Textiles and Information and Broadcasting, criticized Rahul Gandhi, President of the Indian National Congress, for boosting his Twitter following and popularity with fake accounts from Russia and Uzbekistan. Irani accused Gandhi of using bots to divide citizens on caste lines. According to Twitter Audit, 69% of his followers are fake. Amit Shah, who is the President of the Bharatiya Janata Party, and Shashi Tharoor, a Member of Parliament for the Indian National Congress, both have large fake-follower counts, too: 67% and 62%, respectively. Analysis by *The Times of India* indicates that a significant

number of Gandhi's bot followers reveal Russian, Kazakh and Indonesian characteristics, when examining their Twitter feeds. Upon inspection of feeds, the authors found accounts were retweeting incongruous political issues in India and from around the world. On October 15, 2017, Gandhi (@OfficeofRG) retweeted a tweet by President Donald Trump that praised American-Pakistani relations, adding the following text: "Modi ji quick, looks like President Trump needs another hug." It was retweeted over 30,000 times.

Amid the Cambridge Analytica (CA) scandal, both the Bharatiya Janata Party and the Indian National Congress were accused of having contracted CA's parent company Strategic Communications Limited (SCL) for election campaigns. Both parties deny the allegations. According to the *Washington Post*, Christopher Wylie tweeted documents that suggested SCL had conducted behavioral research and polling for at least six state elections between 2003 and 2012 and the 2009 national election. It was unclear in the documents where the data had originated from, whether Facebook or elsewhere, and whether data was misused. Wylie compared SCL's work to "modern day colonialism". According to Praveen Chakravarthy, who works in data analytics for Indian National Congress, CA made a proposal to work with Congress on the 2019 general elections for an "indicative" budget of U.S.$389,460. Reportedly, CA issued a 49-page proposal in August 2017 for the upcoming national election as well as assembly polls in Karnataka, Madhya Pradesh and Chhattisgarh. CA proposed using Facebook data to influence "voter intention." Apparently, the proposal was made by former CEO, Alexander Nix. Chakravarthy stated that Congress did not hire CA's services.

IRAN

Social media came to public attention in Iran in the context of the 2009 Green movement, when the rich and active blogosphere that had for a decade been indirectly political in dealing with societal questions (like social norms, national history, etc.) became explicitly used as a political tool for mobilization against the regime.

The Iranian State (the term is used to refer indistinctly to the government, the military and its branches like the Islamic Revolutionary Guards Corps and paramilitary Basij, which oftentimes differ internally in political objectives and strategies) began systematically monitoring social media activity and criminalizing online activism, as exemplified by the passing of the Computer Crime Law in January 2010 ('Islamic Republic of Iran: Computer Crime Law', 2012 report by Article 19). In September 2009, the Telecommunications Company of Iran had been wholly acquired by the Islamic Revolutionary Guard Corps (henceforth IRGC), strengthening the State's ability to target, trace and block social networking tools, to identify and hunt down opponents,

and to disable cell-phone traffic. Since then, Iran has reportedly been testing a domestic Internet, including national Mehr (the Iranian version of YouTube). It also has actively engaged in cyber warfare under the IRGC, that controls the agenda of the Supreme Cyber Council, established in 2012. Hacking divisions are commonly called 'kittens' and have engaged in cyber attacks, as well as cyber espionage, by setting up personas on platforms like Facebook and LinkedIn; but such activity, due to its military rather than propagandistic aim, is beyond the scope of this report.

News source evidence on efforts at computational propaganda appears to be fairly limited and inconsistent, and while the academic literature is more thorough, opinions diverge significantly on organizational capacity and control. State efforts at computational propaganda in Iran appear to generally respond to two objectives: countering Western narratives and "enemy cultural invasion" in a "soft war" of culture and ideology on the one hand and fostering pro-State and anti-opposition content online on the other, with tactics that range from hacking to social media manipulation. According to Abbas Milani, Iran has an unofficial "cyber army" of 10,000 people dedicated to "cyber-fighting against enemy cultural invasions," although the actual numbers and coordination by the military or the IRGC is controversial. This "cyber army" is held to have hacked opposition and exiled dissidents' Twitter accounts and opposition websites to redirect some users to alternative websites (Arrington, 2009) and to have created a range of websites and blogs which praise the government and are critical of the opposition. It is also alleged that they diffused fake videos showing people aligned to the Green movement burning portraits of founder of the Islamic Republic, Ayatollah Khomeini, in the attempt to incite domestic anger against the protestors, setting up sites that ridicule and mock the 2009 protests, and fake websites that aim at discrediting Western media outlets like BBC Persian, for instance, which is unaffiliated to the BBC but which adopts a similar design to spread ridiculous headlines, in the aim of provoking a sense of nationalism against Western media outlets. Finally, they engaged in viral social media campaigns, as after the 2014 Paris terror attack when Ayatollah Khamenei wrote a letter addressed to the Western Youth defending a peaceful vision of Islam distinct from that of the ISIS. The letter was actively spread under the hashtag #Letter4U on Facebook, Instagram, Twitter, Google+ and Tumblr with links and short messages such as: "Searching for the truth? Then #Letter4u is what you might want to read first," or "Do you know the leader of iran have written a letter for you??".

The 2017 elections, which opposed reformist Rouhani to conservative Raeisi became the object of specific attempts of manipulation by political forces on the three most popular social media channels in Iran: Twitter, Instagram and Telegram ('#IranVotes 2017' report by the Small Media Foundation). The report highlighted the presence of botnets and sock-puppet accounts on

Twitter. The botnet is attributed to the People's Mujahedin of Iran (MEK) revolutionary movement (they claim to be the Iranian government in exile and advocate in favor of the violent overthrow of the Iranian regime) who engage in hijacking hashtags and flooding the platform with tweets decrying the human rights record of the Islamic Republic while exalting the MEK's leader Maryam Rajavi. The sock-puppet accounts were created specifically for the purpose of the election: 84% were identified to be Raeisi supporters, suggesting that 'sock-puppeting' was more prevalent within the conservative camp than it was among Rouhani supporters. On Instagram, both reformist and conservative accounts have produced a lot of content: pro-Rouhani @nedayeeslahat appears to be largely automated, posting 52 times between 13 and 17 May (seven posts within 40 seconds at one point!). A pro-Raeisi sock-puppet account, which was also deactivated after the conclusion of the campaign, is reported to have been directed by a human user. Although Telegram appears to have been less affected, despite its growing importance as the first 'mass online media' of Iran, conservative activists have deployed a fake Rouhani bot with a very similar handle to its official account in an attempt to sway soft Rouhani supporters by spreading anti-Rouhani content including cartoons, news from conservative news' agencies, Qur'anic citations and hadiths, sports news and miscellaneous apolitical memes.

Iran has also been targeted by foreign attempts at social media manipulation: Efandiari reports that, at the height of the 2009 protests, a tweet alleging that Iranian "police helicopters were pouring acid and boiling water on protesters" spread rapidly but was later discovered to be completely untrue (Esfandiari, 2010). Additionally, a closer analysis of the #Iranelection hashtag has since revealed that a vast majority of the tweets containing the hashtag actually originated outside Iran (ibid.). Interesting allegations against Iran that detail its efforts at online propaganda have been leveraged by UAE-based Saudi channel al-Arabiya. The obvious bias of the source does not allow one to draw any conclusion on the actual efforts deployed by Iran, but it is interesting to note once again that accusations of social media manipulation have been used to discredit enemy states. Finally, there is a controversial case of propaganda aimed at an English-speaking audience which remains unattributed. The campaign attempts to leverage hashtags related to the U.S. government (#FBI, #CIA), Israel, Saudi Arabia, and some conservative U.S. hashtags but its motivations remain a mystery, especially given the blatant failure of such an unsubtle campaign. Hypotheses, in addition to a surprisingly badly orchestrated campaign, concern rivalries within the different bodies in charge of social media propaganda in Iran attempting to discredit one another with the failure of the campaign, or foreign attempts at exposing Iranian computational propaganda activities.

ISRAEL

Israel's computational efforts fall within the broader effort of public diplomacy (in Hebrew, *hasbara*) which has in the last years grown more professionalized and centralized in character (Aouragh, 2016). The first reports trace back to 2008 when the Israeli Defence Forces (henceforth IDF) set up its YouTube channel, and efforts have since then involved a series of organizations more or less loosely affiliated with the Israeli State. The general aims of public diplomacy efforts are addressed at both domestic and foreign audiences and include fostering pro-Israeli narratives and countering BDS (boycott, divestment and sanctions) propaganda online that threatens to delegitimize Israel.

The IDF has an Interactive Media Division (of a few dozen soldiers) in charge of spreading Israel's point of view on social media. Their tasks include translating messages, creating graphics and video materials, and coordinating a talkback team. They occasionally rely on volunteers, especially students, such as for the 2012 Pillar of Defence operation, which included 1,600 student volunteers at the Interdisciplinary Center (IDC) Herzliya, a private university. Those involved posted pro-Israel messages online without identifying themselves as affiliated to the government and were offered in return full or partial scholarships. Although mainly student-led, the initiative was overseen by the prime minister's office and was publicly praised by Prime Minister Benjamin Netanyahu. It received U.S.$778,000 in funding (which a New Zealand news magazine *Jacobin* claims came from the American pro-Israel lobby groups Israeli-American Council and Maccabee Task Force) and was meant to operate in parallel with the government's public diplomacy effort, which purchased posts for more visibility on social media. This pattern was repeated in a series of operations, such as Israel Under Fire, which is staffed by 400 volunteer students and disseminates anti-BDS tweets. The Ministry of Foreign Affairs and other institutions (pro-Israel think tanks, and the advocacy groups Reut Institute and StandWithUs) have additionally run public diplomacy hackathons, which include international volunteers.

Members of the personnel for state-led efforts are often recruited as part of military service, which is compulsory in Israel (three years for men, two for women), after which they remain available to return to the army as part of their reserve duty. Coming out of the army, many find work in Israeli software companies which then constitute a great part of the "start-up nation" resources and, in particular, in the very large Israeli communications-related technology industry, which is second only to that of the United States, with a yield of U.S.$5 billion in exports in 2016.

Other coordinated efforts arise from the Ministry of Strategic Affairs, currently led by ex-IDF intelligence official Gilad Erdan. The Ministry allocated more than U.S.$100m in support of

"hidden propaganda" against the BDS movement and its sympathizers. The biggest expenditure (U.S.$740,000) was reportedly budgeted to promote content on social media and search engines, including Google, Twitter, Facebook, and Instagram, while U.S.$570,000 was spent on building Act.il, an anti-BDS app on which supporters were encouraged to spread content online by enrolling on "daily missions" to advance pro-Israel messaging on social media. These missions consist in liking and commenting specific tweets, Facebook posts or petitions with ready-made content or links to videos and cartoons. The public diplomacy effort also enlisted newspapers (like the Yedioth Group) to publish articles and interviews in print and online, aimed at fostering pro-Israeli sentiment without disclosing financing (U.S.$100,000) both in Israel and abroad. The online branch of the group, called Ynet, published promotional videos produced by the Ministry of Strategic Affairs, as well as three paid-for interviews with a ranking official at the ministry. Finally, some funds were channeled to proxy organizations outside Israel, as part of a network of pro-Israeli think tanks and associations.

Unit 8200 (a large unit that is part of the Intelligence Corps and is also responsible for developing communications-related technology like hacking, encrypting and decoding information, at times compared to Britain's GCHQ), with Arabic-speaking agents, is known for monitoring social media life in the Arab world and in particular Palestinian civilian social media activity, following which arrests have ensued. They were exposed for having engaged in practices such as revealing the sexual preferences of civilians so as to blackmail them and engage them as collaborators, or exploiting vulnerabilities such as economic hardship or the need for medical healthcare in Israel for similar purposes, as exposed by a whistle-blowing letter signed by 43 serving and former 8,200 reservists.

Finally, in addition to this coordinated network of computational propaganda around the Israeli State, political forces like the prime minister's Likud party are held to have paid sock-puppets and trolls to plant fake comments online praising Likud members and denigrating the party's rivals, as reported by Haaretz, in the context of the 2015 presidential election.

ITALY

Social media manipulation in Italy is repeatedly described as an "ecosystem," coordinating different types of initiatives mostly affiliated with populist forces such as the Lega Nord (Northern League) and the Movimento Cinque Stelle (M5S, 5 Stars Movement). Public concern has arisen specifically in two crucial political moments: the 2017 Constitutional Referendum and the general elections of 4 March, 2018, in which M5S came first.

For the elections, the Northern League in particular has mobilized "voluntary" Twitter bots among its follower base, retweeting the party's official Twitter feed, @LegaSalvini, as well as with an app that automatically embeds party postings in supporters' timelines. The M5S has for a long time been affiliated with a series of blogs, "independent news" outlets and social accounts that often share misleading or alarmist stories about tragic events and hyperpartisan pieces about immigration, echoing nationalist and Islamophobic rhetoric, and conspiracy theories in the run-up to the general elections. Some of those sources are directly controlled by the party leadership, such as TzeTze (1.2 million followers on Facebook), which pushes content from a network of sites, owned by co-founder of the party's firm Casaleggio e Associati. There are links to a secretive Italian Catholic association La Luce di Maria (Mary's Light) whose Facebook page has 1.7 million followers, which reportedly controls Web 365, a massive network of websites owned by Giancarlo Colono, including 175 domain names and controlling a network of Facebook pages from health to betting odds, which includes two of Italy's biggest Facebook news pages: DirettaNews and iNews24. Finally, websites affiliated both to the Lega Nord and M5S have been shown to share IP addresses, Google Analytics and AdSense IDs, shared also with pro-Putin propaganda site iostoconputin.info, alien disinformation site nonsiamosoli.info ("we are not alone"), euro-skeptical eurocrazia.info, conspiracy complottisti.com, etc. The Reuters report 'Measuring the reach of "fake news" and online disinformation in Europe', however, relativized the impact of such sites in Italy, both in terms of average monthly reach and time on those websites, although Facebook interaction exceeded that produced by most popular news brands.

These pages have targeted politicians from the Partito Democratico (PD, Democratic Party in power before the last elections): TzeTze described Prime Minister Matteo Renzi as a dictator, a liar, a usurer, and a pimp. Ex-president of the Chamber of Deputies Laura Boldrini (also PD) has been the victim of violent online abuse threatening her with rape and murder. Unattributed fake news claimed that her sister was in charge of 340 cooperatives taking care of assistance to migrants (her sister died 20 years ago) and accused her brother of unduly benefiting from political nominations. This last was shared by the parliamentary head of the Lega Nord, politician Gian Marco Centinaio. Maria-Elena Boschi, PD deputy and minister, was also accused of having attended the funeral of mafia boss Toto Riina, while the picture shared actually came from the funeral of a Nigerian immigrant, as shared by page Virus5Stelle which in turn manages pages like M5S News.

The Italian government has taken official action against fake news, instituting educational initiatives in schools and a police unit within the Polizia Postale (Postal and Communications Police), encouraging cooperation between ISPs (including platforms, and Facebook in

particular), citizens and police to report fake news, leading up to public refutations and removal requests.

KENYA

The presence of cyber troops activity in Kenya was brought to international attention during the 2018 Cambridge Analytica revelations by Christopher Wylie. Wylie's leaks confirm Privacy International's reporting that Uhuru Kenyatta's Jubilee Party paid Cambridge Analytica U.S.$6 million for the August 2017 national election, which Kenyatta won with 54% of the vote against his opponent Raila Odinga of the NASA Party (44.9%). After the results were annulled and a new election was rescheduled to take place in October 2017, Odinga withdrew from the presidential race. The NASA Party was previously the Orange Democratic Movement (ODM).

In 2013, approximately four million Kenyans were using Facebook. In 2017, estimates vary between five and seven million Facebook users and ten million WhatsApp users. Mobile penetration is at 87% and Internet speeds are among the fastest in the world. In 2017, over 19 million Kenyans registered to vote. *The Nation* suggests the recent memory of the 2007 violent elections and relatively low presence of Kenyans on Facebook for computational propaganda curtailed the potential impact of misinformation campaigns. The Jubilee senator Kipchumba Murkomen told Reuters that social media campaigns have limited efficacy in Kenya. He stated, "Kenya is not America. In Kenya, vernacular radio stations are more influential than those things." By contrast, GeoPoll research has found that around 49% of Kenyans use social media as their primary source of election news. It is projected that many more, young, technologically-adept Kenyans will be voting for the first time in the next election in 2022.

The 2017 election was characterized as the Kenyan election most affected by fake news. According to a GeoPoll survey conducted in May 2017, 90% of Kenyans reported they had encountered false information regarding the vote, 87% of which reported the information as deliberately false. Social media consistently ranked lower than mainstream media on trust, however. GeoPoll found that Facebook and WhatsApp are the most popular social media platforms for news, preferred by 46% and 25%, respectively. On Twitter, content was spread with two core hashtags: #ElectionsKE and #ElectionsKE2017. Reportedly both parties used bots and fake accounts on Facebook, with Jubilee hiring Cambridge Analytica and NASA hiring Aristotle Inc for data analytics. News websites, including Foreign Policy Journal (fp-news.com) and CNN Channel 1 (cnnchannel1.com), were set up to spread fake news during the election. The sites' branding resembles official international media outlets.

Cambridge Analytica has worked with Jubilee since the 2013 election. Kenyatta won this election in a context of International Criminal Court investigations into crimes against humanity in the 2007 election and in 2018 post-election. Francis Muthaura, Uhuru Kenyatta and Mohammed Ali were charged with five counts of crimes against humanity, including murder and persecution. The court subsequently withdrew the charges against Kenyatta due to insufficient evidence. The Jubilee Party presented its politicians as the victims of a Western plot to persecute Kenyan politicians. Christopher Wylie's predecessor, who worked on Kenyatta's 2013 campaign, was found dead in a Kenyan hotel room.

The managing director of Cambridge Analytica, Mark Turnbull, was caught in a series of undercover videos by Channel 4 News (a British news outlet) discussing extortion, fake news and political branding. He stated: "We have rebranded the entire [Jubilee] party twice, written their manifesto, done two rounds of 50,000 surveys." Turnbull states that Cambridge Analytica wrote all speeches and staged the entire campaign. Nanjala Nyabola of Aljazeera News comments on the irony of the party's vocal criticism of neo-colonial interference while paying a British firm U.S.$6 million to spread said criticism. Cambridge Analytica described its work in 2013 as the "largest political research project ever conducted in East Africa," conducted "based on the electorate's real needs and fears," where the former is jobs and the latter is tribal violence.

During the 2017 election, fake accounts and political advertisements targeted Kenyans on social media. Targeted content followed profile sampling of 47,000 people by Cambridge Analytica. *The Nation* newspaper reported ethno-nationalist rhetoric in targeted advertisements and search engine optimization so that when Raila Odinga was searched on Google, Internet users would be presented with results leading to websites either supporting Kenyatta or criticizing Odinga. Misinformation spread about ethnic violence, where footage and images from historical violence were presented as current violence. The 'Real Raila' video campaign, posted by an anonymous group, presented a dystopian future, where Odinga was portrayed as a "lord of war" and "lord of poverty", who revokes the Constitution and dissolves Parliament by 2020. On YouTube, the video has been viewed over 142,000 times and 480,000 times on Facebook. Texas-based Harris Media was identified by Privacy International as the source of the video. A number of disinformation videos were spread on WhatsApp, reports *The Nation*. Odinga commented: "If it [data manipulation] is allowed to succeed there will be no point of having elections. People will be like robots that vote according to how they are manipulated. That is not the meaning of democracy."

The *Washington Post* has reported that bots were active on Twitter, calling reports of protests in local and international news "fake news." In addition, Twitter bot activity was identified as

being involved in discrediting David Ndii, an economist critical of the government, by trending the #DavidNdiiExposed hashtag. Reportedly, the campaign was not successful. Furthermore, Twitter users in Kenya have responded critically to "#JohoBots", bots promoting Mombasa and the Mombasa County Governor. The reception by users criticized the #JohoBots distraction from political issues.

Fake news was circulated in print, for instance in April 2017 a fake front page of the *Daily Nation* was circulated in Busia County during the primaries. It claimed the local ODM politician, Dr. Otumo, had defected to Jubilee. Its intention had been to discredit Dr. Otumo on the day of nomination. The Busia Country is dominated by the OSM party and hence the fake news were circulated to undermine Otumo's candidacy. The fake front page resembled the branding and design of the *Daily Nation*, giving it false credibility.

Independently, a developer, George Waweru, has built polling bots for each country for the Telegram app. The bots offered polls for candidates for the presidential, senatorial, and member of parliament positions, among others. According to a report in June 2017, 1,000 users were interacting with the bots.

MALAYSIA

Fake news has been around in Malaysia for many years. On the one hand, opposition groups use the term 'fake news' to describe regime propaganda. The regime, on the other hand, has used it to counter questions and critiques posed by local and international news portals. A new portal, Sebenarnya (the truth), was set up in March 2017 by the Malaysian Communications and Multimedia Commission (MCMC), purportedly to enable Malaysians to check the validity of news (Nain, 2017).

Freedom House reported that both government and opposition figures are known to pay online commentators or cyber troops to generate favorable content and slander their opponents (Freedom House, 2017). For example, local news outlets reported that the party Parti Rakyat Sarawak (PRS) was banking on its cyber troops to help the party defend its six parliamentary seats in the 2013 general elections (Yap, 2012). Also, "in January 2017, the ruling party UMNO urged all its members to master the use of the social media to win the war of perception ahead of the 2018 elections. In March, the party called on local divisions to activate newly formed IT bureaus to "counter the slander" on social media" (Freedom House, 2017).

Weeks before 2018 Malaysia elections, bots flooded Twitter with tens of thousands of pro-government and anti-opposition messages. According to Reuters, many of the graphics attached to the tweets credited UMNO's information technology department and some provided details of social media pages of BN-linked accounts. However, Reuters was unable to establish where the tweets originated or which firm or individual may be behind the bot accounts (Ananthalakshmi, 2018).

The Reuters report also found evidence of foreign interference in online political debates in 2018. According to the investigation, nine of the top 10 most active bot accounts containing anti-opposition hashtags and pro-government messages had Russian-sounding names and used the Cyrillic script. Donara Barojan, a research associate of the Washington-based Atlantic Council think-tank, told Reuters that: "The prevalence of bots with Cyrillic screen names does not suggest that Russian social media users are meddling in the Malaysian elections, but does indicate that whoever is behind the campaign purchased some bots created by Russian-speaking bot herders." (Ananthalakshmi, 2018).

Also, there is evidence that CA Political, an offshoot of Cambridge Analytica, supported Malaysia's Barisan Nasional coalition in Kedah state during the 2013 general election, with "a targeted messaging campaign highlighting their improvements since 2008," according to a statement on CA Political's website (Boyd, 2018)

According to local news sources, in the run-up to the 2013 election, the PRS used cyber troops to gain political support. The party relied on its Unit Media Baru (UMB), a five-member team comprising party members, selected based on their interests and wide-ranging knowledge about political issues and happenings in the state. These members allegedly underwent training in Kuala Lumpur together with other UMBs from other component parties prior to the 2011 state election (Yap, 2012). UMB had been entrusted with the task of countering allegations and slanderous statements against the party's coalition on the Internet and to give the general public "the true picture" of what was happening in the country (Yap, 2012).

More recently, several news outlets have reported what appears to be coordinated regional bot action across many Southeast Asia countries, in which Twitter has been flooded by large numbers of anonymous new account holders (Reed, 2018). Though Sri Lanka and Malaysia are the only two countries whose bots have been studied by researchers thus far, a surge in anonymous accounts have been spotted in Cambodia, Vietnam, Myanmar, Thailand, Hong Kong, and China (Seiff, 2018). Thousands of bot-like accounts have followed prominent users.

These accounts use generic names, have no profile photo, no bio and no tweets, which might be evidence of a new "bot farm" (Russell, 2018).

Just weeks before the May 2018 elections, Reuters reported that bots were flooding Twitter with tens of thousands of pro-government and anti-opposition messages. The information technology bureau of the UMNO, which was the ruling party until then, said it was not behind the bots and it did not know who was (Reuters, 2018).

Blocking has also being used as a tool to curb online political activity. According to Freedom House, several news websites (both national and international outlets) were blocked in 2015 and 2016 for reporting on a billion-dollar corruption scandal implicating former prime minister Najib Razak, including the publishing platform Medium (Freedom House, 2017). For example, the popular website Malaysian Insider was banned in February 2016 after publishing a controversial report about the OMDB scandal (BBC News, 2016). Additionally, the MCMC periodically instructs websites to remove content, including some perceived as critical of the government (Freedom House, 2017).

A researcher at the Digital Forensic Research (DFR) Lab of the Washington-based Atlantic Council think tank said more than 44,000 pro-government and anti-opposition messages were tweeted by upwards of 17,000 bots in the weeks before the 2018 elections, held in May (Ananthalakshmi, 2018). A source close to the matter said Twitter had suspended 500 accounts involved in the messages on the Malaysian election since they involved spam or malicious automation (Ananthalakshmi, 2018).

According to Reuters, the tweets included visuals illustrating Malaysian government policies and questioning the opposition's promises:

> "The tweets also include hashtags: either BN's campaign slogans or anti-opposition phrases or both. The hashtags that express disapproval of the opposition coalition Pakatan Harapan (PH) include '#SayNoToPH' and '#KalahkanPakatan', which means "Defeat Pakatan" in Malay. Two of the anti-opposition hashtags - '#SayNoToPH' and '#KalahkanPakatan' - were used around 44,100 times by 17,600 users during April 12-20 and 98% of the users appear to be bots." (Ananthalakshmi, 2018).

In a recent event related to the 2018 elections, Joe Lee, a social media consultant, launched the social media campaign #pualangmengundi, or "go home to vote," which aimed at connecting

voters too poor to afford plane and bus tickets home with sponsors stepping in to fund their travel. The hashtag reached trending topics within hours, but then it was "hijacked" by bots, which overwhelmed the timeline and disrupted attempts to match sponsors with voters. According to Lee, the bots were flooding the timeline with thousands and thousands of pro-government messages (Seiff, 2018).

MEXICO

News outlets have reported that all three dominant political parties in Mexico have used social media to manipulate public opinion in various elections at the national and state level. Such interference in the public debate has come both directly from government officials and party personnel, and also indirectly from private actors hired by them.

Regarding political actors, many outlets, such as the MIT Technology Review, have presented evidence that the PRI used bots in the presidential campaign of 2012 (Orcutt, 2012). Regarding private actors, local papers and TV stations have long run flattering or condemnatory stories in exchange for purchase of advertisements or money. More recently, as platforms such as Facebook, Twitter, and Google have become important sources of information for many Mexicans, they are increasingly being used as channels to spread misinformation (Orcutt, 2012). Also, many outlets have reported the operation of political marketing companies in Mexico specialized in the provision of digital services. Services offered by such private actors include bots management, cyber attacks, trend creation, crisis control, and creation and dissemination of fake news (disinformation, spoofing or manipulation) (Uriel, 2017).

Mexico was also mentioned in Channel 4 News revelations about the Facebook scandal involving Cambridge Analytica, in April 2018. Forbes published information mentioning the existence of ties between the PRI and Cambridge Analytica, suggesting a "modus operandi" similar to the one in the United States (Forbes Staff, 2018).

Reports have also called attention to the interference of foreign actors and foreign governments in Mexican politics. In a speech in December 2017, Lieutenant General H.R. McMaster, then the U.S. national security adviser, said there were "initial signs" that the Russian government was trying to influence the 2018 Mexican elections, but provided no further details (Peinado, Palomo, & Galán, 2018; Semple & Franco, 2018). In 2018, a digital strategist told *The Sydney Morning Herald* that he identified 4.8 million items about López Obrador that had been posted to social media and news websites in the past month by users outside Mexico. About 63% were associated with users in Russia, he said, and 20% with Ukraine. But other digital consultants,

running different programs, said they had found no such evidence of election-related social media activity from Russia, and Cossío's findings could not be independently verified (Semple & Franco, 2018).

Many different tools and techniques have been used to disseminate political propaganda. Based on the content analysis of news articles, three main strategies were identified: (1) the use of bots and trolls to disseminate political propaganda, (2) the creation and dissemination of fake news, (3) hacking and surveillance of citizens by the government.

There is much evidence that bots and fake accounts have been used by political campaigns in Mexico for many years. These tools are used to influence voter behavior by spreading false stories and also in order to promote or to attack other candidates. One strategy is to coordinate automatic and simultaneous tweets with specific words and phrases to turn a message or topic into a "trending" topic on Twitter (Peinado et al., 2018).

According to a BBC News article, all major Mexican political parties have been involved in social media manipulation for years—the earliest reported activity is from 2010. Different sources have reported that, during Mexico's presidential elections of 2012, bots were used to support political campaigns. In particular, there is evidence linking the Institutional Revolutionary Party (PRI) to tens of thousands of bots (Martinez, 2018; Orcutt, 2012; Soloff, 2017).

In 2017, Tanya O'Carroll, a technology and human rights adviser for Amnesty International, published an investigation of the political impact of bots and trolls in Mexico (O'Carroll, 2017). An article by the BBC describes a video showing the operation of a "troll farm" in Mexico, where people were tweeting in support of Enrique Peña Nieto of the PRI in 2012 (Martinez, 2018).

According to a report published by *El País*, the main target of parties' online strategies are young people, including 14 million new voters who are expected to play a decisive role in the outcome of the July 2018 election (Peinado et al., 2018). Thus, one of the strategies employed by these bots was the use of profile photos of attractive people from other countries (Soloff, 2017).

Evidence of the use of bots was also found at the local level. Alberto Escorcia, a Mexican reporter who specializes in network analysis, reported that bots were used in the 2017 governorship race in the central state of Mexico—the country's most populous state. According to him, messages spread through automated accounts against candidate Delfina Gomez of the National Regeneration Movement party (Morena) helped elect PRI candidate, Alfredo del Mazo. Escorcia

says the messages spread by bots against Gomez were mainly attacks based on her gender (Martinez, 2018).

Bots and trolls are also widely used in Mexico to spread fake news. Parties have been generating fake news and disseminating them both on social media platforms (either using sponsoring tools or bots) and on websites that pretend to be media (Gutiérrez Rentería, 2017). Experts and fact-check reporters say that a new technique in the 2018 campaign entails sending out messages through WhatsApp or Snapchat (Martinez, 2018; Peinado et al., 2018). Rumors and misleading hashtags are also often disseminated to undermine social protests online (Freedom House, 2017).

In June 2017 *The New York Times* exposed the use by the Mexican government of Israeli security software known as Pegasus to hack into the cell phones of political opponents and civil society activists (O'Neil, 2017). Investigations showed that human rights defenders, journalists and anti-corruption activists have been targeted by the spyware, which the Mexican government allegedly purchased to be used only to investigate criminals and terrorists (Ahmed & Perlroth, 2018).

A Mexican web developer, Iván Santiesteban, reported that around 20,000 bots were used in the month and a half before election day 2012 to create online conversations favorable to Peña Nieto (PRI). About 1,000 people worked with the Peña Nieto campaign that year to "combat negative comments in social media and instead position the positive ones," the chief of marketing for the campaign, Aurelio Nuño, has said (Semple & Franco, 2018).

According to *El País*, Mexico's political trolls in Mexico are often young college students who need the extra income and earn about 12,000 pesos a month (€520). Each troll was in charge of dozens of fake Twitter or Facebook accounts that used either fictitious or stolen identities (Peinado et al., 2018).

Andrés Sepúlveda, a Colombian hacker, told in an interview with Bloomberg Businessweek in 2016 how he rigged political campaigns throughout Latin America from 2005 to 2015. In the Mexican presidential campaign of 2012, he received U.S.$600,000 to manage an army of 30,000 bots to work in favor of Peña Nieto. The services included stealing campaign strategies, manipulating social media to create false waves of enthusiasm and derision, and installing spyware in opposition offices (Robertson, Riley, & Willis, 2016).

Victory Lab is a digital marketing agency in Mexico City. The 2015 municipal elections in the State of Mexico was the first time the agency used its sites to spread false news. Its CEO told the magazine *Expansión* that the news pages operated by the agency on Facebook had an average of 500,000 fans. According to him, political parties and candidates sought the agency to hire false news strategies and bots on Twitter and Facebook, especially during the debates prior to the election (Chávez, 2017).

Other websites reported that complete media strategy services can cost more than one million pesos (U.S.$50,000), while the monthly cost to create Twitter trends range from 20–30,000 pesos (U.S.$1,000 to $1,500) (Uriel, 2017). Benito Rodríguez is the leader of the collective @100tifika, which runs 133,000 members, dedicated to tweeting and creating hashtags. According to Rodríguez, a work scheme only on Twitter has an estimated cost of at least 17,000 pesos for 28,000 tweets, which can convert a hashtag into a national trending topic (Chávez, 2017).

In an event in November 2017, when Mexican President Enrique Peña Nieto introduced José Antonio Meade as the candidate to succeed him, hundreds of fake accounts tweeted out @JoseAMEade until the Institutional Revolutionary Party (PRI) candidate's name became a trending topic, according to Alberto Escorcia, a Mexican reporter who specializes in network analysis (Peinado et al., 2018). In March 2018, fact-check agency Verificado flagged as fake news a report that said an opinion poll commissioned by *The New York Times* showed Meade leading the presidential race with 42% of voters behind him.

Also in 2018, a message circulated on social media announcing that most Mexicans would have to re-register within days if they wanted to vote in the presidential election. It was a piece of fake news that set off a low-level panic on Facebook, Twitter and other platforms (Semple & Franco, 2018). Other examples of fake news that have circulated in Mexico include a claim that Pope Francis weighed in on the presidential race and criticized López Obrador's political ideology (Semple & Franco, 2018).

In another episode, a Twitter hashtag (#YaMeCanse25), used to protest against the disappearance and murder of 43 students in 2014, became the most used hashtag in Mexican history—and immediately came under attack by bots. According to experts and activists, automated accounts started tweeting the hashtag out with links to pornography or violent photos in the hopes that Twitter would flag the hashtag as spam and block it, however: "Twitter activists responded by adding numbers to the end of the hashtag, changing the number every time they came under attack. They eventually got to #YaMeCanse25. Digital activists in Mexico

often refer to the agents that hijacked these hashtags as "Peñabots" since they defend Peña Nieto and the political establishment" (Soloff, 2017).

MYANMAR

Amid many ethnic disputes in Myanmar, U.N. investigators have accused Facebook of playing a "determining role" in violence against the Muslim minority living in the country. There is evidence that the social media platform has been used by military-led campaigns to coordinate acts of violence against this community (Barron, 2018). Fake news claiming that Muslim worshippers are attacking Buddhist sites, for example, are quickly shared and amplified on social media and this misinformation has been connected with a surge in anti-Muslim protests and attacks on local Muslim groups (Frenkel, 2016). Much of this content has been blamed on the nationalist monk organization *Ma Ba Tha* (Baker, 2016).

Frenkel argues that social media has been used to mobilize radical Buddhist anti-Muslim groups and gather supporters not only across the country, but also internationally—connecting these groups with extreme movements around the world, including the more radical, nationalist American groups, like the Ku Klux Klan, that have supported Trump (Frenkel, 2016).

According to Freedom House, "regional government officials, elements of the military, and some business people are believed to hire cyber troops to promote their cause and spin the news on social media—especially Facebook—to their advantage, though their activities have not been well-documented." (Freedom House, 2017).

Digital researcher Raymond Serrato reported having found evidence of a connection between social media activity and military operations against the Muslim Rohingya. According to Serrato:

> "A Facebook group associated with a Buddhist nationalist organization known as *Ma Ba Tha* appears to have started posting in June 2016, and accelerated its activity the following October when an insurgent ambush triggered brutal army reprisals. Leading up to a second wave of attacks in August 2017, the number of posts again exploded with a 200% increase in interactions." Serrato also scraped data from a military Facebook page which revealed similarly timed activity spikes (Barron, 2018).

Various hate speech monitoring organizations told *The Myanmar Times* that nationalist individuals and groups are exploiting a lack of digital literacy to "dehumanize" minorities and "encourage violence" online. The monitoring groups described a trend of anti-religious and anti-

ethnic material in the guise of fabricated news items and "information posts" on social media (Baker, 2016).

According to a spokesperson at the Myanmar ICT for Development Organization (MIDO), fake accounts are being set up to spread hate and encourage violence against Muslims. The aim is not only to shield the original authors but also to create a larger echo chamber around these topics. A number of posts were framed like news articles with headlines and accompanying pictures, and several used videos. For example, one Facebook post, which repurposed real footage from a 2013 sectarian conflict to claim that riots against Muslims had just broken out in the streets of Mandalay, was shared 16,598 times (Baker, 2016).

In May 2017, Myanmar's government warned the public that false news and rumors saying President Htin Kyaw would step down were being spread by unidentified people wishing to cause "political instability." State-run media reported that accounts with false names were used to spread the rumors and named two Facebook accounts that it said had published "fabricated news." (Reuters Staff, 2017).

More recently, journalists in Myanmar and Vietnam have reported dozens or hundreds of new followers on Twitter that they suspect to be bots, often with names common in their respective countries, since late March—a trend that was also reported in other Southeast Asia countries (O'Byrne, 2018; Reed, 2018). Many of these new Twitter accounts do not have any activity yet, use repeat variations of the same names, or adopt pseudonyms such as "Myanmar Egg;" and there are suspicions among those affected that they are either "bots" designed to generate automated posts, or people acting in concert while trying to disguise their identities (Reed, 2018).

David Madden, founder of Yangon-based tech hub Phandeeyar, said he had attracted more than 1,000 "fake followers" over March 2018. Raymond Serrato, a social media analyst affiliated with Democracy Report International, has studied some of the suspicious new accounts and said that the accounts appeared to be choosing who to follow based on keywords in users' profiles, and then automatically following the first accounts Twitter suggested (Reed, 2018).

News outlets also reported evidence of surveillance. In March 2018, Myanmar's government said it would spend more than 6.4bn kyats (U.S.$4.8m) to establish a system for monitoring social media (Reed, 2018).

NETHERLANDS

Media coverage in the Netherlands of misinformation online largely concerns Geert Wilders, the leader of the far-right People's Freedom Party (PVV), in his campaign for the 2017 national elections and in the upcoming 2018 municipal elections. Wilders has earned a nickname as the 'Dutch Trump' and was recognized as part of a pan-European far-right, anti-Islamic movement. In the March 2017 national elections, Wilders' party became the second largest party in the lower chamber of the Dutch parliament with 13.1% of the vote, after Mark Rutte's conservative People's Party for Freedom and Democracy (VVD), which gained 21.3% of the vote. While the international perception of Wilders' politics were understood as part of the wider European populist movement, *The Guardian* has noted that the domestic perception of Wilders' prospects were not as significant as otherwise portrayed. The Dutch media covered a variety of political contest, including new populist far-right parties like the FvD and VNL, popular cosmopolitan parties D66 and GL, a 'Turkish' party called DENK and the social democratic party, the PvdA.

The Guardian's analysis encourages one to be careful of exaggeration of automation when commenting on automation in Wilders' 2017 campaign. One source comments on Wilders' use of the hashtag #Kominverzet, Dutch for #computation, which is automatically retweeted by bots. Mentionmapp Analytics Inc, a social media data analytics firm, tracked the #Geertwilders hashtag in February 2017 and found 26 fake accounts which systematically amplified the #Geertwilders hashtag. The accounts were found to have similar "fingerprints"—that is, date joined, follower-to-following ratio and tweets-to-like ratio. The accounts all retweeted one tweet by @abermans (the account and tweet have been removed). The *Financial Times* has extensively reported on Wilders' Twitter activity. Wilders discredited the validity of mainstream media, for instance by sharing repeatedly a cartoon of President Trump pitching a ball labeled 'social media' over a crowd of reporters into the hands of an average person sitting in an armchair. Wilders' follower count dwarfs his party's account by a 330:1 ratio on Twitter. By contrast, the ratio of Rutte's followers to his party's followers on Twitter was 1.5:1.

The *Financial Times* found spikes in new followers for Wilders following particular events, such as his conviction for race-related discrimination offenses on December 9, 2016 and his comments on the Berlin Christmas market terrorist attack on December 19, 2016. By taking a random sample of 100,000 users that follow Wilders and Rutte's prime ministerial account (@MinPres), FT Data analyzed the frequencies of the followers that mentioned the accounts of RT, Sputnik and the top five Dutch news outlets over a six-month period. The FT found that Wilders' followers were 12 times more likely to mention Sputnik and nearly eight times more likely to mention RT than Rutte's followers.

Ahead of the 2018 municipal elections, the PVV have produced a 2.5 minute anti-Islam video, broadcast on TV and the Internet. The video claims Islam represents violence, persecution of Christians, death penalty to apostates, forced marriage, among other things, before dripping blood down the screen. The public prosecution department in the Netherlands will not take legal action against the video because it has been identified as not encouraging hatred, discrimination or violence because it targets the Islamic religion, and not Muslims. Wilders has welcomed the decision, tweeting: "We will continue to show the facts about Islam soon, on television, at the time of the political parties' broadcast." It is too early to comment on a wider mis- or disinformation campaign during this election; this propaganda video, however, provides an indication of the tone of misinformation in the elections.

NIGERIA

In Nigeria, approximately 25.7% of the population had Internet access in 2016. According to Freedom House's 2017 'Free the Net' report, compared to traditional news media, online media is relatively free from restrictions, with no blocking or filtering of online content. Social media, such as YouTube, Facebook, WhatsApp and Twitter, are available and among the most popular websites in the country. Among the barriers to Internet connectivity in Nigeria are gender imbalances and languages. The Web Foundation and Paradigm Initiative found that in Lagos, Nigeria's largest city with approximately 21 million people, women were 50% less likely to have Internet access than men of their same age, education and income. Freedom House reports that, while most Internet content is in English, there are over 500 local languages in Nigeria. A variety of factors have impeded Internet growth, which is a crucial context for understanding the reach and efficacy of cyber troops in Nigeria. The 2015 Cybercrime Act has been used to arrest online content producers who post content critical of the government. The 2015 Frivolous Petitions Prohibition Bill, which punished critical expression on social media, was withdrawn in 2016.

The 2018 Cambridge Analytica leaks revealed that Cambridge Analytica's parent company, SCL Elections, has interfered in two Nigerian elections. Christopher Wylie posited that SCL Elections had helped with a "rumour campaign" in the 2007 election, spreading fears that the "election could be rigged." Anti-election rallies were reportedly organized, with the persuasion and support of religious leaders, to dissuade opposition supporters from voting in 2007. The 2018 leaks contended that a Nigerian oil billionaire hired SCL Elections in December 2014 to work on sitting President Goodluck Jonathan's campaign in the 2015 election in order to keep him in power. Jonathan was contested by Muhammadu Buhari, who had hired David Axelrod of AKPD, a former strategist for President Obama, to support his digital campaign. It was revealed that 16 million Nigerian Facebook users' data had been mined by SCL Elections and used to spread

targeted fear and misinformation. Reportedly, SCL Elections were paid U.S.$2.8 million to "orchestrate a ferocious campaign" against Buhari.

The 2015 misinformation campaign involved spreading anti-Islamic videos on social media, threatening that Buhari would enforce Sharia law. The anti-Buhari videos sought to portray Buhari's Islamic faith as a violent prospect for the nation, with content showing people being dismembered and dying in ditches. The aim of these videos was to incite fear in Nigerian voters. In addition, rumors were spread about Buhari's education and health, sparking speculation about whether he was fit to run for office. Reportedly, SCL Elections hired an Israeli intelligence gathering firm, Black Cube, to hack Buhari's emails. Black Cube has denied the allegations as a "flagrant lie" and, while SCL Elections confirms it was hired for "advertising and marketing services," it denied the use of hacked information. The next government elections will take place in 2019, which according to observers are at risk of more computational propaganda.

Abdullahi Tasiu Abubakar of City University of London has reported on the Nigerian army's cyber troops offensive against Boko Haram. The army has confirmed it using "scientific measures" to find anti-government content online and to remove Boko Haram videos from social media. Abubakar has reported on the removal of YouTube videos and the blocking of social media accounts in an effort to disrupt "the online activity of the jihadists." Additionally, according to a former army media consultant, the security services use psyops in their counter-insurgency campaign. Dr Abubakar commented: "The extent to which the Army has engaged in PsyOps is difficult to know and that is the nature of PsyOps. They have never publicly explained why they engaged in psychological operations – nor have they even admitted using them."

PAKISTAN

In Pakistan, news sources reported that fake news has become especially problematic in Pakistan, with all leading political parties asking their social media teams to create fake profiles as part of their social media strategy (Shahid, 2018). There have been suspicions that parties from both left- and right-wings are relying on bots to propagate their messages. In fact, Pakistan was one of the countries mentioned by Mark Zuckerberg, when he declared concern about the risk of Facebook being used to manipulate elections in 2018 and announced that the platform will require all political advertisements to clearly mention who is paying for the message and for their identity to be verified.

According to news site The Diplomat, an anonymous social media executive of the ruling Pakistan Muslim League-Nawaz (PML-N) reported that almost "everyone is running fake

Facebook accounts and Twitter bots" in Pakistan to keep "pace with what others are doing" (Sohail, 2018). According to *Pakistan Today*, the opposition has accused the ruling party of using the IT Ministry to manipulate the 2018 election. There are suspicions that PML-N could be purchasing user data through its power over the IT Ministry. On the other hand, the opposition party Pakistan Tehrik-e-Insaf (PTI) is accused of using trolls/cyber trolls/bots to create hashtag trends in Twitter, in order to face-off criticism on social media (Sohail, 2018).

Some evidence of foreign interference was also mentioned. According to a local newspaper, the Chairman Joint Chief of Staff Committee (CJCSC) General Zubair Mahmood Hayat has declared that Pakistan's enemies were spreading rumors about the country, acting on a methodically planned agenda (Daily Times, 2018; Pakistan Today, 2018).

There were also suspicions of activities in Pakistan affecting politics in India. India's Union Minister Jitendra Singh declared in May 2018 that "studios" have been set up in Pakistan "under a well-planned strategy" to promote propaganda and manage content on social media to mislead the people of Kashmir. The minister of state said a "perception is created in these studios on a daily basis and false messages are spread" (Press Trust of India, 2018).

A common strategy in Pakistan is the use of fake news to discredit political actors online, often by accusing them of blasphemy, a criminal offense which carries a death penalty (Freedom House, 2017). Fake news has been on television for many years, but more recently it is also spread through social media posts and WhatsApp messages. The use of WhatsApp to spread disinformation has increased over the last two years. The messages are usually short, meant to be consumed quickly, and sent on to as many people as possible, usually with the added phrase "forwarded as received." Many of them are analyses of the political situation of the country (for example, rumors that the PML-N had hired Cambridge Analytica's services for the upcoming elections; Shahid, 2018), or rumors about the wealth of politicians and leaders. Disinformation is also targeting political and human rights activists, including rumors about activists committing alleged blasphemy or working on "foreign" agendas (Shah, 2018).

Freedom House has identified no documented examples of cyber troops paid to distort the online landscape (Freedom House, 2017), however some local news sources have reported a number of Twitter trends that are either sponsored or built using automated help, raising suspicions of bot activity (Sohail, 2018). *Pakistan Today* has reported that, with the nearing of 2018 general elections in Pakistan, hundreds or even thousands of Twitter bots have been identified, allegedly sponsored by political parties, which are allocating resources to

manufacture consent and develop favorable opinion in order to gain maximum votes (Sohail, 2018).

According to The Diplomat, social media managers from the ruling Pakistan Muslim League-Nawaz (PML-N), and the two main opposition parties PTI and the Pakistan People's Party (PPP) have declared, off the record, that "creation of fake Facebook and Twitter accounts to propagate their narratives was the official policy of each party." (Shahid, 2018). Kaleem Hafeez, a member of the PTI social media team, told The Diplomat that his party is not ruling out the possibility of the PML-N purchasing data to manipulate elections, considering the party's control over the IT ministry (Shahid, 2018).

Other reported techniques used by Pakistani authorities to curb political discourse are censorship and blocking (technical filtering) to limit access to political, religious, and social content online. In June 2017, the Berkman-Klein's Internet Monitor reported that Pakistan "blocks news and human rights websites and content critical of the faith of Islam," as well as sex and nudity, and tools used to circumvent censorship or protect privacy (Freedom House, 2017).

A very popular case of fake news became international news and turned into an international relations issue. In 2017, the then defense minister, Khawaja Muhammad Asif, wrote a Twitter post directed at Israel after a false report—which the minister apparently believed—that Israel had threatened Pakistan with nuclear weapons (Goldman, 2017). The Pakistani defense minister said that "Pakistan is a nuclear state too" after a fake story saying that Israel had threatened to destroy Pakistan. The story appeared on 20 December on the site AWD News, which has been identified by fact-checking organizations as a fake-news site (Graham-Harrison, 2016). The fake story about Israel even misidentified the country's defense minister, attributing quotations to a former minister, Moshe Yaalon. Israel's current minister of defense is Avigdor Lieberman. The Israeli Defence Ministry responded on Twitter to say the report was fictitious (Goldman, 2017).

One interesting case of political manipulation involved a Twitter hack. Someone hijacked a former American NFL player's Twitter account and started tweeting about Pakistani politics. One tweet, responding to a tweet by the prime minister's daughter concerning the upcoming 'Panama Papers' inquiry, was retweeted more than 1,700 times, including by a number of local journalists. Because the Twitter account was verified, people thought it had greater legitimacy (Brandom, 2017).

SERBIA

Most efforts around computational propaganda in Serbia seem to be concentrated around the ruling Serbia Progressive Party (SPP) and its leader Aleksandard Vucic who became prime minister in 2014 and was elected president in 2017. The party has recruited a team of trolls, with contingents in every town, working under a veil of secrecy, who are hired as civil servants (the monthly salary, according to Deutsche Welle's informant is €370). There are about a hundred individuals in this team, who manage thousands of identities to comment on news outlets, stifling opposition campaigns by comparing them to Western operatives or praising the government.

According to a series of leaks from 2015 published by the web portal Teleprompter.rs and summarized by the SHARE lab report, in 2014 and 2015, the ruling party SNS has been using different types of software that could be used for astroturfing and other means of public opinion manipulation by relying on volunteers. There is a special "Internet team" within the party, which includes public officials like councilpersons at the City of Belgrade, or positions at the Office for Media of the President of the Republic. The first program, named Valter, consisted of software that activists from the SPP downloaded on their devices, and which was subsequently turned into a bot that would give positive or negative comments on Serbian news outlets. The second version, named SkyNet, also allowed monitoring of comments published, ranking activists themselves according to results. The third and final version of the program, Fortress, relied on Facebook and could also be used for DDoS attacks. This resulted in uncommon patterns of activity on news media, with some comments receiving more than 5,000 votes (and up to 50,000 in one case!), disproportionate to the website traffic for the websites. It is important, however, to qualify these findings: only half of the population uses the Internet, while the most used media in Serbia remains the television, which is consumed 5 hours a day on average, and if governmental propaganda succeeds it is in great part due to the government's influence on broadcasting.

Governmental control over social media is tight: in the context of the 2014 floods, when the government response was absent or uncoordinated, citizens organized aid and volunteers through social media. Several, however, were called in by the police for questioning and threatened with charges such as "spreading panic" for their posts. In the meantime, governmental tabloids spread fake click-bait news about hundreds of dead bodies floating on rivers or Roma gangs on the loose.

Several unofficial campaigns were launched against Serbia's ex-province and most recent independent state, Kosovo, in particular in the context of the country's bid to access UNESCO

membership. The campaign relied on volunteer "internet warriors" and sought to emulate Israeli tactics online. Marko Djuric, the head of the Serbian government's office for Kosovo, announced a campaign in 2015 to "spread the truth" about the poor living conditions of the Serbian minority in Kosovo, a discourse which is often picked up by Russian outlets like Sputnik, which are quick to denounce violence against Serbs in Kosovo, continuing their support of the Serbs, who have not generally received good coverage in Western media since the war.

Interestingly, this discourse in Western media outlets is often tied to fears concerning strengthened Russian presence in the Balkans: MEP McAllister, interviewed by *The Guardian*, expressed concerns about Russia's influence over Serbian anti-Western, anti-EU propaganda. In contrast to this position, the prestigious Italian geopolitical magazine *LIMES*, has in another article analyzed alarmism about Russian interference in the Balkans as part of a broader wave of alarmism about Russian propaganda, concluding that the presence of narratives similar to those actively circulated by Russian media actually spread without direct manipulation (see the case study report on Hungary above).

SOUTH AFRICA

According to a survey of 33,000 people by the Edelman Trust Barometer (reported by News 24), trust in news and government are at staggering lows: 14% trust the government, 82% think the government is a broken institution, 62% cannot separate fake news from real news and 54% say it is becoming harder to know if news is credible or not. These are challenges for the new president, as of February 15, 2018, Cyril Ramaphosa. Admittedly, distrust in media and government has swept across a number of liberal democracies.

South Africa enjoys an Internet penetration rate of 54% (ITU 2016) and relatively little government-led manipulation online. In 2017, David Mahlobo, the Minister of State Security, responded to journalists' questions on fake news in South Africa. Mahlobo stated his department had considered regulating social media in order to contain fake news, but he acknowledged that a solution would not be easy and could be seen as "interfering with human rights."

The most significant event of social media amplification involved the billionaire Gupta brothers, who were accused of corruption and meddling with Zuma's government. It was revealed that the Guptas paid the British PR agency Bell Pottinger to improve their media image amid a watchdog investigation into their corruption. Work with Bell Pottinger reportedly started in January 2016. *The Guardian* reported that Bell Pottinger was paid £100,000 per month by

Oakbay Investments, which is owned by the Guptas. Another source reported that Bell Pottinger were paid U.S.$2 million for the whole operation.

The Bell Pottinger strategy sought to divert attention from the Guptas' influence on government by attacking "white monopoly capital," which was shared by Gupta-controlled or Gupta-aligned actors on social media. "White monopoly capital" content attacked white-owned businesses for disproportionately influencing the government. In particular, it sought to stir up anger about race-related "economic apartheid." The campaign hired a Florida-based digital marketing firm, Devumi, which deployed botnets to amplify the "white monopoly capital" phrase. According to Dr. Crofton Black of the Bureau of Investigative Journalism, many of the accounts were clone accounts whose feeds revealed incongruous automation where political causes from a wealth of countries were being retweeted. Black highlights that this case reveals the global reach of automation and amplification in digital politics on social media, orchestrated by private actors who can be geographically far removed from the political environment where their services are deployed.

Black and Fielding-Smith at the Bureau of Investigative Journalism identified Devumi as the source of automation in the Guptas' social media campaign through an unlikely source—Eric Klien, founder of Lifeboat Foundation, an organization dedicated to future planning given the existential threats of AI. Klien had used Twitterboost to boost his content, which is also known as Devumi. Black and Fielding-Smith collected 18,000 tweets containing "white monopoly capital" between July 12 and August 22, 2017. Ten core accounts were found to push this phrase systematically: they tweeted this phrase over a hundred times during this period. A second tier of 200 accounts tweeted the phrase consistently, between ten and sixty times. Some 6,000 accounts tweeted the phrase once. Many of the accounts posed as South Africans, but were in fact run from India and promoted fake-news websites like wmcleaks.com, wmcscams.com, whitemonopolyafrica.com and whitemonopoly.com. Black commented in an interview that it was clear that the accounts were bought in a big batch and the attribution to Devumi was owed to Klien. Black emphasized one unanswered question, however: while Bell Pottinger bore the burden for stirring racial tensions in South Africa, their involvement with the bots is still unknown.

According to Professor Steven Friedman of the University of Johannesburg, the Guptas' astroturfing campaign sought to convey that the mostly white private sector was just as guilty of state capture. Friedman commented to CNET: "What it was supposed to do is to make the campaign against the Guptas appear to be racist, and because 90 percent of South Africans are

black, it was assumed that there would be a huge well of black support, and that didn't really happen."

The Facebook pages of Weekly Xpose (alternate news, owned by the Guptas), Black Opinion (owned by Black First Land First), WMS Leaks (a fake-news site), ANN7 (a 24 hours news channel owned by the Guptas) and *The New Age* newspaper (owned by the Guptas) posted content supporting Zuma and the Guptas. On Twitter, between July 2016 and June 2017, "white monopoly capital" was retweeted 215,000 times. In July 2017, Black Opinion was taken down for inciting racial hatred by writing articles on white monopoly capital. It was restored two weeks later. Fake accounts were used to trend #RespectGuptas.

Twitter has shut down around 900 bots that systematically spread white monopoly capital content and promoted the above mentioned fake-news websites in the form of botnets and sock puppets. Bot accounts, such as Khumbu Malan, and fake-news accounts, such as @WMC–LEAKS, were removed from Twitter for publishing defamatory stories about journalists like Peter Bruce, the former *Business Day* editor. The WMC Leaks website remained up and, according to *Business Day*, a new Twitter account called WMCLEAKS was created immediately after the former's shutdown. One story claimed that Bruce was paid by political elites and "white monopoly capital" in order to discredit his journalism.

The investigators' analysis of 5,000 retweets of "white monopoly capital" indicated that many accounts, such as Pigs and Pints (@PlgsandPints) and LBDT (@VaieenCaarp), were clones of already existing accounts belonging to real people, but with altered usernames. These fake accounts played an important role in the astroturfing campaign. The bureau's investigators found that the bots were not effective, however. On the contrary, accounts—whether automated or human-administered—that tweeted in favor of the Guptas were ridiculed as "Guptabots." According to the *Daily Maverick*, many fake accounts tweeted in a style of English that was foreign to South Africans, making their detection often straightforward. Bots would not tweet in other South African languages, even when engaged in those languages.

The *Daily Maverick's* network analysis of Guptabots revealed activity from within South Africa, too: the "BLF" node in the network— BLF standing for 'Black First Land First'. Accounts in this network tended to retweet anything by BLF spokespeople, blogs like blackopinion.co.za and defamatory memes. The authors contend that, due to the content distributed by BLF supporters, it is difficult to differentiate between fake accounts and real people.

SOUTH KOREA

South Korea is one of the most affluent countries in Asia and a pioneer in high-speed and wireless Internet; nearly every household is connected to the Internet (BBC, 2018). Like many aspects of South Korean politics, South Korean social media manipulation is shaped by the country's complicated relationship with its Communist North. South Korean intelligence services often justify deploying 'psychological warfare' on the grounds that it is needed to defend against Northern propaganda (Sang-Hun, 2013). Democratic activists have long feared that these capabilities will be exploited by the intelligence services to intervene in domestic politics.

These fears have been confirmed in recent years, as a developing scandal revealed that employees from the National Intelligence Service (NIS) had launched smear campaigns using fake accounts against South Korean opposition parties in the lead-up to the 2012 presidential election (The Korea Herald, 2013). The NIS wrote more than 5,000 posts critical of North Korea since 2009; many of these also accused opposition parties of sympathizing with North Korea (Sang Hung, 2013). A civilian whistle-blower also said that he was paid U.S.$450–540 per month for posting pro-government comments on various web forums between 2008 and 2009 (Freedom House, 2017). In 2014, the country's former intelligence chief Won Sei-hoon was convicted of trying to influence the presidential election. However, Won was in charge of a very small team of only nine agents, which suggests this was not a systematic government strategy (Benedictus, 2016).

SYRIA

The Assad family—members of the Syrian Ba'ath Party—have been in power in Syria since Hafez al-Assad seized power in the 1970 military coup. Although the Ba'ath party is a secular Pan-Arab organization, the minority Alawite elite has come to dominate both the party and the military, becoming increasingly repressive as opposition to their leadership grew (Economist, 2000). The rule of Hafez's son, Bashar al-Assad was challenged in 2011 by the Arab Spring protests; following violent repression by the government, these protests transformed into a complex and brutal war involving both regional and international actors (BBC, 2018). The seven-year-old civil war in Syria has been described as the "first social media war" and the "the first skirmish in the Information War," (O'Neill, 2013; Diresta, 2018).

Syria's Internet infrastructure is severely damaged, highly decentralized and often subject to significant censorship; the Assad regime has long attempted to assert total control on political communication (Freedom on the Net, 2017). However, more than 200 media workers have been

killed since the start of the revolt, which means that both Syrians and international audiences have increasingly come to rely on social media for information (BBC, 2018). YouTube in particular has become crucial as activists use cell-phone video to document the human toll of the conflict, leading to projects such as *The New York Time*'s effort to sort, verify, and contextualize videos coming from the conflict (Stack, 2018). Each side in the dispute has waged its own propaganda offensive: the radical Islamist group Islamic State (also known as ISIS, ISIL or Daesh) is widely recognized as a successful innovator in this field (Berger and Morgan, 2015), and both regional powers—such as Iran—and global powers—such as the U.S. and Russia—have been accused of spreading computational propaganda in the conflict (Di Giovanni, 2016; Cockburn, 2016). However, this article will focus on social media manipulation undertaken by or linked to the Syrian government.

An early report on Syria's cyber troops argued that the government invested in Twitter bots to overwhelm the pro-revolution narrative with pro-government posts (York, 2011). York also noted that the government had outsourced to a Bahraini company, EGHNA, which has been successful in flooding the #Syria hahtag. EGHNA's usual project cost is about U.S.$4,000 (EGHNA, 2017). Another prominent actor is the Syrian Electronic Army (SEA), a hacker group which is widely considered to be supported by the Syrian government (Harding and Arthur 2013; Stork, 2014). In a 2011 speech in Damascus, Assad likened anonymous online warriors to his frontline troops: "There is the electronic army, which has been a real army in virtual reality" (Harding and Arthur 2013). The SEA combines cyber attacks and propaganda, for example by using phishing to take over the social media accounts of Western news outlets (Harding and Arthur, 2013). In 2013, the SEA hacked the official AP Twitter account and tweeted that Barack Obama had been injured in an explosion, which tipped the stock market by U.S.$136 billion (Fisher, 2013). Pro-Assad activists reportedly earn around U.S.$500–$1,000 for high-profile attacks on Western targets (Harding and Arthur, 2013). Harding and Arthur (2013) argue that these attacks serve the double purpose of punishing Western news organizations critical of Syria's regime and spreading Damascus's alternative narrative.

A recent BBC investigation raised doubts about some of the most influential accounts which back the Syrian government, some of which do not seem to be linked to real persons (BBC, 2018). Recent reports have also emphasized the close alignment between Syrian and Russian propaganda (Palma, 2018; Diresta; 2018). Scott Lucas, Professor of International Studies at the University of Birmingham, has suggested that: "although Moscow became militarily involved in the Syrian conflict in 2015, they had a propaganda office at the presidential palace in Damascus since the beginning." (Palma, 2018). Conspiracies which are promoted by the suspicious accounts identified in the BBC study, such as the idea that the Syrian chemical attacks are a hoax

created by rescue workers known as the "White Helmets," are often widely shared by Russian state-run media outlets such as RT and by Western far-right activists. Similar actors will share similar narratives; many of the same accounts which claim that American victims of mass shootings are actually actors in a "staged" tragedy repeat the same allegation about Syrian war victims (Palma, 2018).

RUSSIA

Reports of Russian social media manipulation and computational propaganda have dominated the international news cycle in the past two years. However, these policies need to be understood both in the context of the Russian presidency's recent consolidation of power and a broader historical background. Vladimir Putin has been the dominant figure in Russia's political landscape since his election as president in 2000; he was re-elected for a third term in 2018 (BBC, 2018). Throughout his presidency, Putin has restricted the independence of various state institutions and the media; this has been accompanied by increasing nationalism and hostility to the West. Russian TV is dominated by channels that are either run directly by the state or owned by companies closely linked to the Kremlin (BBC, 2018).

Within the Russian government narrative, however, this increasing control is primarily a reaction to Western interference in Russia's domestic politics throughout the Cold War and the 1990s. In particular, the "colour revolutions" in Ukraine, Georgia and Kyrgyzstan are seen as examples of foreign meddling (Streltsov, 2011). Many Russians find it improbable that Western politicians brand Russia Today (RT) as a 'foreign agent' or 'Moscow's propaganda arm' while designating media outlets like Radio Free Europe/Radio Liberty as more reliable than domestic news sources (RT, 2017; Reuters, 2017). Russia has advanced its understanding of 'global information warfare' in international forums such as the United Nations and the Shanghai Cooperation Organization, where it has proposed a wider understanding of cyber security which encompasses the dissemination of information "harmful to the spiritual, moral and cultural spheres of other states." (Gjelten, 2010; Franke, 2015).

Russian efforts to fight this information war have centered on the Internet Research Agency (IRA). The IRA was founded in 2013 and went on to hire hundreds of employees to set up fake accounts and post pro-Putin, anti-Western content online, with a particular focus on targeting Ukraine and other Eastern European countries (Elliot, 2014; Graff, 2018). It first came to Western attention following Adrien Chen's 2015 article 'The Agency' in *The New York Times*. The IRA attracted young professionals looking for "simple, well-paid work" by paying higher than average salaries—about U.S.$700 a month, according to former workers who have been

interviewed by Western media (Graff, 2018; Wigham, 2018). It was run similarly to any other marketing agency, with departments focused on graphics, data analysis, and search engine optimization, as well as IT and financing (Barrett, 2018). Estimates of its total staff differ widely, from 400 to 1,000 (Graff, 2018).

By 2014, the IRA had begun to expand to targeting the U.S. population through YouTube, Facebook and Twitter, with the stated goal to "spread distrust toward the candidates and the political system in general." (Graff, 2018). IRA employees also traveled to the U.S. "to collect intelligence for their interference operations," according to the indictment. The pages set up to target Americans—with titles like "Secured Borders," "Blackivist" and "Heart of Texas"—together amassed hundreds of thousands of followers (Graff, 2018). The IRA also used stolen social security numbers and fake and stolen identity documents in order to establish 'sock puppets' or fake identities. They used fraudulent bank accounts to purchase political advertisements—taking advantage of the capacity of many online platforms for micro-targeted messaging. The team harnessed bots to amplify hashtags like #Trump2016, #TrumpTrain, #MAGA, and #Hillary4Prison. The hashtags, advertisements and images shared predominantly opposed presidential candidate Hillary Clinton and supported Donald Trump (Shane, 2017; Shane and Goel, 2017). IRA instructions stated: "use any opportunity to criticize Hillary and the rest (except Sanders and Trump—we support them)" (Graff, 2018). Lastly, the IRA team escalated to organizing political rallies, for example in New York and Washington D.C. By the fall of 2016, the IRA was spending upward of U.S.$1.25 million a month on influencing the U.S. election alone.

As Graff (2018) summarizes: "The sheer volume of the IRA's effort staggers the imagination. All told, it posted some 80,000 pieces of content in 2015, 2016, and 2017. Facebook has struggled to wrap its arms around the IRA's activities in the year since the election; according to Facebook's estimates, more than 126 million Americans saw some of the IRA's propaganda. The company estimates that the IRA spent around U.S.$100,000 to promote some 3,000 different advertisements related to the campaign, all part of what it says are about 470 "inauthentic accounts and Pages." On Twitter, the Russian efforts garnered more than 131,000 tweets, and there were more than 1,000 videos uploaded to YouTube."

In February 2018, the U.S. special counsel investigation into Russia's interference in the 2016 U.S. election led by Robert Mueller indicted 13 Russian nationals and three organizations for "conspiracy" to illegally influence the U.S. presidential campaign. Although news stories have largely focused on the IRA, the Mueller indictment also revealed details about a network of

affiliates which funded the IRA, many of which were connected to Yevgeny Prigozhin, a wealthy Russian oligarch closely connected to Putin (Graff, 2018).

Despite the indictment, Russian bots have been accused of interfering with a variety of U.S. political issues in 2018: supporting Republicans on energy policy, and amplifying conspiracy theories around the Parkland shooting (Frenkel and Wakabayashi, 2018). Others have cast doubt on these findings, suggesting that Russian trolls have become an excuse for any negative outcome (Ingram, 2018; RT, 2018). Crucially, however, such accusations are not limited to a U.S. context, even if these tend to draw the most attention. Similar methods have been used to systematically undermine and divide Ukrainians and promote Assad in Syria (Diresta, 2018; see also *Ukraine* and *Syria* in this report.) Snegovaya (2017) details a sophisticated campaign to target the Russian minority in the 2017 German elections. A *Guardian* investigation found that in the U.K., tweets from members of the IRA "troll army" were quoted more than 80 times across British-read media outlets (Hern, Duncan, and Bengtsson, 2017). Recent information warfare skirmishes have raged over the attempted assassination of the Skripals in the U.K., which British officials have linked to Russia. British official estimate at least 2,800 bot accounts have posted messages or retweets about the Skripals, reaching at least 7.5 million people in Britain (Liam, 2018).

The worldview of the Russian government, in which it is under constant attack from Western information dominance, has led Russia to pioneer some of the most innovative and sophisticated computational propaganda. Many other countries—both targets and allies—have consciously or unconsciously begun to imitate these techniques (Diresta, 2018).

THAILAND

Freedom House's report found no public documentation of paid actors manipulating political content on the Internet until 2017, though there were organized efforts to restrict political engagement online coming from the military government, which has been the main challenge for freedom of expression in Thailand. Officials offered financial incentives to citizens to monitor one another online and in some cases created fake accounts in order to join secret chat groups—even baiting users to criticize the monarchy or the *junta* (Freedom House, 2017).

There were vague reports of fake news in the Thai media landscape. Some outlets reported that as social media become more popular, false information is spreading. In particular, the popularity of messaging apps like the Japanese LINE and WeChat makes it easier for individual users to share messages and photos, but also to pass around false information. For example, a

video widely shared on LINE claims that a sandwich maker substituted pork with cotton. The news articles, however, focused on government efforts to fight back (Rojwanichkun, 2017). In one recent episode, a Cambodian man has reportedly been arrested in Phnom Penh after allegedly posting fake news about the Thai prime minister on the Internet while six Thais have been detained in Bangkok for sharing it (Bangkok Post, 2018).

Censorship has been more often used to curb political discourse. Since the launch of the coup, the *junta* has suppressed freedom of expression with a number of laws, including the Computer Crimes Act, which was amended in 2017 to make penalties even harsher for spreading false information. However, increased censorship and control over the flow of information is not a new phenomenon in Thailand. Freedom of speech and media was especially restricted during the 30-day period following King Rama IX's passing on October 13, 2016. According to Freedom House, the government enforced a mourning period for all media outlets and requested that ISPs cooperate to monitor online content for blocking or deletion. "More than 1,370 websites were shut down in October alone, according to the Associated Press. December 2016 profile published on the BBC Thai website became famous overnight for pulling no punches, and was shared over 2,000 times on Facebook. It was quickly blocked." Like blocking and filtering, content removal also increased after the death of the king (Freedom House, 2017).

There are records indicating that Thai government agencies also possess surveillance technologies. Freedom House reported that spyware from the Milan-based Hacking Team was bought between 2012 and 2014, and that Thailand has also obtained licenses to export telecommunications interception equipment from Switzerland and the U.K. (Freedom House, 2017).

Like other Southeast Asia countries, bot activity has recently been reported in Thailand. Since the beginning of 2018, thousands of newly created Twitter accounts have been following Thai online influencers including journalists, media companies, scholars and celebs. All are new accounts, have no followers and, in almost all cases, no tweets. They have authentic Thai-sounding names such as @Fah12113 or @Thanaphorn_1230. Some user names are written in Thai script, but all of those have machine-generated strings such as @hjZuotIwLtiSojc and @hIrQMl1B71tIYKF as account names. Few have profile photos, and those that do look like any face plucked from the Thai social mediaverse (Ruiz & Saksornchai, 2018).

The arrival of an "online ghost army" has been reported in several Southeast Asia countries, but they are all very locally aware. In each country, the identities use regionally authentic names, languages and profile photos to follow local influencers. Because the accounts have no activity

so far, it is still uncertain whether they are machine- or human-made and what are their intended purpose. They could be used for commercial enterprise, state actor, organized crime or rogue algorithm (Ruiz & Saksornchai, 2018).

Regarding the "bot army", some news outlets have reported that about 400 new fake followers were identified by one media broadcast. In another case, the online newspaper Khaosod English usually adds about 200 Twitter followers per month, but more than triple that number —697 and counting—appeared in early 2018 (Ruiz & Saksornchai, 2018).

In early 2017, according to the *Bangkok Post*, Thai police and soldiers raided a rented home yesterday near the Cambodian border, discovering an alleged "clickfarm" run by three Chinese nationals. The house had hundreds of mobile phones wired to computer monitors and Thai SIM cards. Officers originally thought the men were running a fraudulent call center, but the suspects said they were being paid to operate a vast network of bot accounts on WeChat, China's largest social network. "According to the Post, the trio of men said a Chinese company (which they refused to name) supplied the phones and was paying them each 150,000 baht per month (about U.S.$4,403) to artificially boost engagement on WeChat for products sold online in China. The operation was reportedly headquartered in Thailand due to the country's relatively cheap smartphone usage fees." (Deahl, 2017).

TURKEY

Starting with the Gezi park protests in 2013, the Turkish government, led by the AKP (Justice and Development Party) ruling party, intensified its presence online, both by exerting tighter control on the Internet by restricting access to the Internet and to social media, by monitoring online content and activity, and by coordinating efforts at computational propaganda by means of AKP-affiliated bots and trolls. In the same year, the pro-AKP *Star* reported that the AKP was launching "a massive project to boost the party's social media presence by hiring over 6,000 new employees for its newly formed social media team," with AKP social media representatives in over 900 districts and 1,000 staff to be located in Istanbul, 600 in Ankara, and approximately 400 in İzmir, the three main cities of the country.

Pro-government coordinated bot campaigns initially fell under the responsibility of the youth branches of the AKP, whose members came together to discuss what they would tweet and when, "but the government quickly realized this wasn't very efficient [while] the people they mobilized had loyalty but they weren't online experts." In September 2013 efforts had professionalized. The AKP recruited its new social media team, known as the New Turkey Digital

Office, which was responsible for converting AKP sentiments into trending hashtags and engaging in abusive behavior against journalists and civil society movements (like the "Vote and Beyond" electoral monitoring campaign) in Turkey. News sources have reported a centralized botnet, playing out with the trending topics in two ways: astroturfing (by first boosting a pro-government hashtag) where possible, or hashtag hijacking. This was visible in the recent example of the oppositional hashtag, #DemirtaşıÇokÖzledik ("We missed Demirtaş a lot") campaigning for the jailed pro-Kurdish Peoples' Democratic Party leader, where the bots boosted counter-messages to demoralize his supporters. In addition to this botnet, prominent columnists and editors are often enlisted in such efforts, and civilian Twitter accounts with big follower bases can suddenly be repurposed, such as in the case of the 2015 elections in which "an account with a 'sexy girl profile picture' suddenly changed its name and brand to launch a smear campaign using its 42,000 followers." against the aforementioned election monitoring group Vote and Beyond.

Social media monitoring increased after the July 2016 coup attempt, as Turkey's General Directorate of Security, the high command of the country's police, officially asked the public to report any social media account that praised the coup or had a "criminal element," and set up hotlines to deal with citizen reports of "terror propaganda." "Terrorism" in the Turkish context is an ambivalent term, as such an accusation often extends to all government critics, who are increasingly prosecuted and imprisoned. In early 2016, 1,656 were arrested for allegedly supporting terrorism and 10,000 investigated.

Individual targeting mainly consists in accusations of being a "traitor," a "terrorist," or a "terrorist supporter." 2,000 cases of online abuse, death threats, threats of physical violence, sexual abuse, smear campaigns and hacking have been reported, as part of an AKP campaign to intimidate journalists. In 2016, trolling was turned into a real lynching, as the *Hurriyet* newspaper building attacked by protestors. There is also a gendered element to such attacks, as female journalists are most often targeted by hundreds of trolls with degrading sexual-related insults (e.g. as happened to Nevsin Mengu who was a popular CNN-Turk anchorwoman forced out of her job). In early 2018, for instance, Turkey cracked down on public expressions of dissent about the ground offensive against the Syrian Kurdish YPG militia, arresting more than 300 people. The story gained traction as a German-Kurdish player was banned from entering the country after sharing a video calling for participation in a rally in the German city of Cologne to protest against Turkey's military offensive into northern Syria's Afrin region.

UKRAINE

Following the events of the 'EuroMaidan' revolution between November 2013 and February 2014, Ukraine became the target of some of the most sophisticated and disturbing computational propaganda recorded to date (Shearlaw, 2015; Polyakova and Boyer, 2018). Social media was crucial both in the coordination of protests against former president Viktor Yanukovych and in the ensuing conflict in which Russia annexed Crimea, a Ukrainian peninsula in the Black Sea inhabited by a Russian-speaking majority and the eastern 'oblasts' of Donetsk and Luhansk, which devolved into a war involving the Ukrainian army and Russian-backed separatist forces (BBC, 2018). As recorded in Computational Propaganda Working Paper No. 2017.9, notable examples of Russia's purposeful disinformation campaign against Ukraine include widespread presentation of the new Ukrainian government as a "fascist coup" as well as the tragedy of flight MH17, following which a variety of conspiracy theories were widely spread online by human accounts as well as a variety of Twitter bots (Zhdanova and Orlova, 2017).

Although worries about Russian propaganda are widespread in the Ukrainian government, the extent to which the government has responded in kind is disputed. In 2014, the government established a "Ministry of Information Policy" to counteract "Russian information aggression" (Grystenko, 2014). The new ministry announced the launch of an "i-army" based at i-army.org where citizens and volunteers can access and share "truthful" information on social media (Benedictus, 2016). Media reports suggested that almost 40,000 people registered as "information soldiers" with the ministry; however, Zhdanova and Orlova (2017) argue that the volunteers did not receive any specific direction and that the government's response more generally has been "sporadic and weak". They argue that Ukraine's response to Russian computational propaganda has been relatively decentralized and largely driven by civil society organizations such as StopFake and the Ukraine Crisis Media Centre.

However, some journalists have also alleged that the Ukraine has seen the rise of "Kremlin-style trolling," i.e. organized online abuse against those critical of government positions (Gorchinskaya, 2017). Although there is no definitive evidence that this is coordinated by the government, recent survey of media professionals showed that 56% of respondents believe in pro-government manipulation in online debate (Internews, 2017). Furthermore, both the ministry and the i-army initiative were widely criticized by journalists and civil society activists who compared it to Orwell's "Ministry of Truth" (Recknagel, 2014). More recently, the most prominent actions taken by the government have been bans on Russian web services and social media networks such as VK, Odnoklassniki, mail.ru and Yandex (BBC, 2017). Crucially, Ukraine's media landscape is in many ways still dominated by TV; therefore the government's policies on

Russian-language television stations are in some senses more important than those related to computational propaganda (Ennis, 2014; Szostek, 2014).

UNITED ARAB EMIRATES

UAE efforts at computational propaganda form part of a coordinated military and public diplomacy effort, and cannot be examined in isolation from other initiatives, like the funding of think tanks and conventional media, in a highly funded campaign to disseminate narratives favorable to the regime, which promotes itself as a role model of "liberal authoritarianism," in opposition to Islamism and Iranian and Qatari expansionism. This well-orchestrated campaign is one of the reasons, according to interviewed expert Andreas Krieg (Assistant Professor at the Defence Studies Department of King's College London and strategic risk consultant working for governmental and commercial clients in the Middle East, including the Qatari government) why there is very little independent research on the matter, as most think tanks focusing on the Middle East and the Gulf States are directly or indirectly funded by the UAE. Although there are very few news sources on the matter, Krieg compared Emirati efforts to the Russian ones, albeit unsuspected, being predominantly under the radar.

Emirati attempts at computational propaganda reportedly started out as part of a more defensive strategy in late 2012, spreading positive messages about the UAE, targeting mainly U.S. and—to a lesser extent—U.K. audiences. In the context of the Arab Spring, however, the UAE deployed a more aggressive strategy including foreign attacks on any type of political Islam, conflating any form of political Islam with Islamic State-type Salafi-jihadism; against Qatar in the first place, as well as Turkey and Iran; but also against opposition forces in Syria, Iranian proxies in Iraq (such as Islamic Revolutionary Guards Corps section outside Iran), Egyptian grassroots movement Tamarod, etc. For this purpose, the UAE reportedly set up and funded local online news outlets tied to social media accounts, reinforced by bots and trolls, which, for instance, have orchestrated discredit campaigns against the Tamarod and the Muslim Brotherhood in Egypt, in support of military forces. These also targeted Libyan revolutionary groups, spreading narratives that equated them to terrorists while building consensus around counter-revolutionary forces such as the national army of Libya, in favor of the Gadaffi regime.

This strategy was further refined in the context of the 2014 Gulf Crisis, especially since advancing national interests online seems to have become a particularly convenient tactic in comparison with more visible kinetic attacks. An army of trolls and bots would disseminate unsubstantiated allegations made by think tanks and experts close to the UAE about Qatar's support for terrorism and Qatar's humanitarian aid to Hamas in the Gaza Strip. The 2017 hack of the email

account of the Emirati ambassador to the United States revealed ties with pro-Israel think tanks in Washington aimed at undermining the image of Qatar.

Since 2014, propaganda efforts have expanded to the West, and campaigns outside the Arab world were outsourced to PR and consulting firms in the U.S., the U.K., Germany, Switzerland, etc., as the Emirates lacked the necessary know-how and capacity. The funding, however, is centralized and reportedly attributable. According to Andreas Krieg who was consulted for this country-profile, Abu Dhabi has created a powerful web of policymakers, think tanks and experts in the United States, aligned to neo-con and AIPAC positions, influencing Washington discourse and positions on Middle Eastern affairs, but more importantly to directly shaping the Trump administration's initial approach to the region. This emerged in the Mueller investigations over UAE attempts at buying influence during the 2016 presidential campaign, but appears to have been played out in person, in the form of visits and lobbying, as business man and adviser to the Abu Dhabi Crown Prince George Nader often visited the White House and met with senior adviser Jared Kushner and former chief-strategist Steve Bannon.

In 2017, when the Qatar News Agency was hacked and published remarks attributing remarks in support of Iran and critical of Donald Trump to the Emir of Qatar, these were further relayed by Emirati and Saudi news channels and spread on social media. This was later proved to have been the deed of Russian hackers on hire with the help of the FBI. The UAE is reportedly responsible for the attack but has denied attribution.

The UAE currently controls a wide web of traditional news outlets, together with their respective social media presence: al-Arabiya (the network is Saudi but operates from Emirati capital Abu Dhabi) which interestingly has been frequently denouncing Iranian and Qatari attempts at computational propaganda, and Sky News Arabic, to cite the two most important. In its endeavor to create a new narrative as a tolerant Middle Eastern partner that shares U.S. security concerns and spreads anti-Qatar messages, the UAE engaged with the U.S. in the creation of the Sawab Center in 2015, which is dedicated to countering Islamic State's online propaganda efforts using social media, as part of the international Working Group on Strategic Communications of the Global Coalition against Daesh. According to the Emirati Minister of State of Foreign Affairs, its aim is to "amplify moderate and tolerant voices from across the region." It has since then released a YouTube video and claims to have reported 500,000 social media accounts that were subsequently shut down and to have launched the social media campaign #deludedfollowers with success. Such efforts, however, are held by Krieg to have been largely ineffective. Their funding, mainly supported by the UAE is in the tens of millions of dollars, and the purpose of the expense seems primarily to be aimed at supporting the UAE in

its efforts to combat radicalization and become the champion of counter-terrorism in the Arab world.

UNITED KINGDOM

A wealth of academic literature and journalistic material has dealt with the presence and extent of computational propaganda in recent elections in the U.K.—including the 2017 general election, the 2016 EU Referendum (or Brexit), and the 2014 Scottish Independence Referendum. Automated content has been reported during the 2017 election. Facebook announced in May 2017 that it had removed 10,0oos of fake accounts to curtail fake news. According to WIRED magazine, an unofficial campaign was conducted on the dating app Tinder. Supporters of the Labour Party amassed over 100 Tinder Premium accounts through donation, a paid subscription which crucially afford (1) unlimited right swiping and (2) location manipulation (where one can select where one would like to be swiping). The supporters deployed bots via the volunteered accounts in contested constituencies and targeted 18 to 25 year olds by swiping right to every individual, thus seeking matches, and upon matching sending an automated political message. The purpose had been to inquire into voting intentions and persuade individuals to vote for the Labour Party. Depending on the response, the bot would reply with, for instance, locations of voting stations or giving its reasons for voting for Labour. One such constituency was Dudley North where, in total, between 30,000 and 40,000 messages were sent. The bots did not identify themselves as automated accounts. Reportedly, the initial budget for this unofficial campaign was £500.

As previous research by the COMPROP group has argued, "During the 2016 UK Brexit referendum it was found that political bots played a small but strategic role shaping Twitter conversations. The family of hashtags associated with the argument for leaving the EU dominated, while less than 1% of sampled accounts generated almost a third of all the messages" [Brexit Data Memo]. 53.6% of content being shared by Twitter users interested in U.K. politics came from professional news organizations; junk news accounted for 11.4% of news content shared on Twitter. Researchers at (1) Swansea University, (2) City University, London, (3) University of Edinburgh and (4) University of Oxford have investigated Russian interference in Brexit. The Swansea research team identified 150,000 bot accounts linked to Russia; while researchers at City found that, of the 794,949 users who had produced 10 million Brexit-related tweets, 37% (30,122) were located in the U.K. Reportedly 16.9% of bot accounts had Russian links (13,493). University of Edinburgh researchers identified 419 Russian bot accounts, while a University of Oxford researcher identified 54 bot accounts. Research by COMPROP found a

large degree of automation on Twitter during the Brexit campaigns but little evidence of Russian links to sources.

The leaks of the involvement of Cambridge Analytica (as well as Aggregate IQ or AIQ and its group holder SLC) in the EU Referendum by whistle-blower Christopher Wylie made public the presence of computational propaganda during the Brexit campaign. The legitimacy of the Brexit vote has been put in question in light of the revelations of the Leave campaign's outmaneuvering of spending limits (by donating £625,000 (U.S.$1 million) to the pro-Brexit student group BeLeave) and the illegality of personal data misuse to target voters. *The Guardian* reported that £3.5 million was spent on AIQ by four Leave campaign groups (Vote Leave, BeLeave, Veterans for Britain, Northern Ireland's Democratic Unionist Party) for targeted political advertising. The Vote Leave campaign spent 40% of its budget on AIQ's services. Meanwhile, *The New Yorker* reported, according to receipts released by the Electoral Commission, that the Vote Leave campaign spent £3.9 million (of its £7 million budget) on AIQ. Wylie stated in his parliamentary hearing that AIQ enjoyed a conversion rate of between 5 and 7% and targeted 5–7 million people during the Brexit campaign. Wylie argued that the four campaigns worked together to outmaneuver the spending limits. The revelations also alleged that Robert Mercer had offered Cambridge Analytica's services for free to Leave.EU; separately, Cambridge Analytica have boasted about working on the Brexit campaign, but claims have since been retracted stating no contract was signed and no work done.

Reports allege that cyber troops were active in the 2014 Scottish Independence Referendum, too. In 2017, it was reported that 388,406 messages were sent by bots during the Independence campaign; in favor of Independence. Automated activity is suspected to have been orchestrated by the Kremlin. In addition, post-referendum where the Remain in the U.K. side won by 55% of the vote, it was reported that fake news spread on Twitter, YouTube and Facebook regarding the interference with the vote in order to ensure a victory for the Remain campaign.

The British government's Intelligence Community, notably GCHQ's Joint Threat Research Intelligence Group (JTRIG), systematically and persistently targets online content by militant Islamic groups, such as ISIS and Daesh, as a part of a wider military campaign. The *Financial Times* reported that Jeremy Fleming, head of GCHQ, stated cyber attacks have "made a significant contribution" to fighting against ISIS. Glenn Greenwald and Andrew Fishman, in their exposé of the JTRIG's activities, detailed that its domestic operations work with the Metropolitan Police in London, the MI5 and Serious Organized Crime Agency among others and their objectives include monitoring "domestic extremist groups," "denying, deterring or dissuading" criminals and "hacktivists" and "deterring, disrupting or degrading online

consumerism of stolen data or child porn." JTRIG has pseudonymously created content and social media accounts as part of a digital strategy designed to "discredit, promote distrust, dissuade, deter, delay, or disrupt" targets.

UNITED STATES

The United States appears to be the case where most attempts at computational propaganda, both governmental and partisan, have been documented; in fact, the adoption of techniques to influence political opinion online seems to have become general electoral and political practice.

The first systematic efforts can be traced back to 2011, when DARPA set up its Social Media in Strategic Communication (SMISC) program, which received U.S.$50 million in funding, with the double aim of both detecting and conducting propaganda campaigns on social media. It financed a variety of studies: some more theoretical (topic trend analysis and sentiment detection, modeling emergent communities and network dynamics), and others more directly linked to online propaganda (automated and crowd-sourced content generation, persuasion campaign structures recognition and effects measurements, as well as counter-messaging tactics). Similar research was undertaken by other branches of the army, with the U.S. Air Force Research Laboratory investigating how human behavior could be manipulated through social networks, or the development of software for the use of "sock puppets" to manipulate social media and influence online conversations.

These operations were required by law to target only foreign audiences, as outlined in the Smith-Mundt Act (2012), which protected domestic audiences from American efforts at "public diplomacy." The Smith-Mundt Modernization Act of 2012 overturned this provision and it is not clear whether this has had consequences in terms of domestic audiences being targeted by computational propaganda efforts, especially since it appears difficult to distinguish between foreign and domestic publics online. Online manipulation seems still to be forbidden to federal agencies, as highlighted by the denunciation by the Government Accountability Office concerning the use of social media tools by government agencies, such as the use of crowd-based "Thunderclaps" to coordinate information spreading, resulting in "covert propaganda." The Environmental Protection Agency, for instance, coordinated a campaign to promote the Clean Water Act rule, which reached 1.8 million people, but failed to disclose the origin of the Thunderclap messages, de facto engaging in astroturfing. The tool was also used, among others, to promote National HIV Testing Day and National Women and Girls Day, amplifying the campaign through tweets, blogs, news updates, letters, and other tactics. Other domestic government-led efforts at influencing opinion were designed to foster support of governmental

policies. In 2014, the government spent U.S.$760 million to hire private advertising firms, according to USASpending.gov, from marketing research to opinion polling to message-crafting assistance, etc.

The *Washington Post* explicitly tied the overturning of the Act to the involvement of the Pentagon in a counter-propaganda initiative against the U.S.-based "extremist" Somalimidnimo.com website, undertaken via U.S. contractor Navanti by means of a messaging campaign that amplified comments posted on the site by readers opposed to al-Qaeda and al-Shabab. Furthermore, a reporter and an editor at *USA Today* were targeted in an online propaganda campaign because of their investigations of Leonie Industries, the Pentagon contractor in charge of info ops in Afghanistan. According to the Post, a minority owner of the firm admitted to having set up the smear campaign, which included the creation of fake websites under the journalists' names, of their Wikipedia pages, posting fake information on forums with the intent of tarnishing the journalists' reputation, as well as Twitter accounts under their names.

Efforts directed at foreign audiences include active campaigns of influence, in line with the U.S. tradition of public diplomacy through print media and broadcasting, as well as counter-propaganda activities. The U.S. Special Operations Command is considered to be the spearhead of propaganda abroad: as first reported in 2008 by *USA Today*, it directs a collection of websites with civilian appearance (including Southeast Europe Times, SES Turkey, Magharebia, Mawtani al-Shorfa and Central Asia Online) known as the Trans Regional Web Initiative, aimed at foreign audiences, that conducts psychological operations to combat violent extremist groups. Deployed to support the military and diplomatic delegations, it now has operational teams in 22 countries. Funding in 2009 peaked at U.S.$580 million a year, but this was progressively reduced to U.S.$202 million by 2014, mostly spent on propaganda in war zones. It has subcontracted to Navanti Group to help conduct "information operations to engage local populations and counter nefarious influences" in Africa and Europe, and to Leonie Industries. The latter, however, was found by the Government Accountability Office in 2013 to have inadequately tracked its operations, whose impact was therefore unclear.

The Center for Strategic Counterterrorism Communications (CSCC) was set up in 2011 to coordinate anti-jihadist and violent extremist campaigns. It managed more than 350 Twitter accounts for the State Department, the Pentagon, the Homeland Security Department and American Foreign allies' accounts in a sock-puppet network. On YouTube, Facebook and Twitter, U.S. diplomats have begun actively trolling ISIS, arguing with pro-ISIS accounts and producing videos portraying ISIS-conquered territory as a hellscape. U.S. military Central

Command coordinated an astroturfing campaign called "Operation Earnest Voice," officially targeting al-Qaeda, the Taliban, and other jihadist groups in the Middle East. It began as a psychological warfare operation in Iraq to combat the online influence of those opposed to the coalition's presence in the country, and is reported to have received more than U.S.$200 million in funding since. The software it developed, contracted for U.S.$2.76 million to Ntrepid, allows to posting from different accounts, covered by a VPN to randomize location and avoid detection.

Finally, computational propaganda has involved agencies outside the Departments of State and Defense: Associated Press revealed in 2014 the U.S.-led initiative to foment anti-government unrest in Cuba by creating a clandestine, Twitter-like service on mobile phone networks called ZunZeo, coordinated by USAid. The network, built with secret shell companies and financed through a foreign bank, lasted more than two years and attracted tens of thousands of subscribers.

Although the use of tools for online manipulation is frequent in political campaigning (Ratkiewicz et al., 2011a; Woolley, 2016), attempts at manipulation appear to have been particularly marked in the context of the 2016 U.S. elections, in both camps, relying on interactive advertisements, live-streamed video, memes, and personalized messaging. According to Woolley and Guibeault (2017), the tacit goal of using these tools was to affect voter turnout, but they were also used to: "achieve other, less conventional, goals: to sow confusion, to give a false impression of online support, to attack and defame the opposition, and to spread illegitimate news reports."

The Democrats relied on astroturfing with "grass-roots tweeters" who were asked to post specific messages and graphics at coordinated, strategic times, such as during presidential debate (a similar strategy was adopted by the Bernie Sanders campaign, coordinated by social media "volunteers" in closed Slack rooms). The Democrats benefited from the support of the Brock network, a large network owned by David Brock which includes the Media Matters for America watchdog website, two pro-Clinton "super PACs," the opposition research outfit American Bridge, the pro-Clinton fact-checking Correct the Record (that, for instance, launched the TrumpLeaks website to "uncover unreported" and unfavorable video or audio of Trump), as well as Shareblue, which with a budget of U.S.$2 million was focused on exposing alleged news coverage against Hillary Clinton and extensively engaged in astroturfing throughout the campaign. On the side of the Republicans, Ted Cruz hired Cambridge Analytica, which then coordinated Donald Trump's campaign, allegedly with its infamous psychographic techniques. The Trump campaign is reported to have similarly engaged in intense astroturfing by means of

viral videos and memes, which behave like political advertisements but which require very little disclosure. Nimble America, a non-profit funded by Silicon Valley millionaire Palmer Luckey orchestrated an anti-Clinton campaign, while the Koch brothers are held to have coached up a "grassroots-roots" army of their own, actually offering online certificate courses in things like "social media best practices" via their conservative advocacy group Americans for Prosperity.

Astroturfing seems to have become commonplace and forms a specialized market: the firm Devumi, for instance, stands accused of stealing real people's identities for at least 55,000 bots out of a network of about two million it possesses. The company's clients covered the political spectrum, from liberal cable pundits to a reporter at the right-wing site Breitbart to political commentator Hilary Rosen and U.S. ironworker turned politician Randy Bryce. Such tactics are also used for specific political issues: North Texans for Natural Gas, a seemingly "grass-roots" group, for example, was discovered to have been funded by four Texas energy companies when it attracted the attention of several media outlets for launching a pro-fracking meme factory.

The growing awareness around the efforts of foreign nations, and especially Russia, to influence U.S. opinion led to the passing of the Countering Foreign Propaganda and Disinformation Act in December 2016 as part of the 2017 National Defense Authorization Act. The provision established a Global Engagement Center that coordinates information sharing across government agencies, to collect and analyze the narratives generated by foreign governments. Of the U.S.$120 million allocated by Congress, nothing has currently been spent and none of the analysts of the agency appear, according to *The New York Times*, to speak Russian. The U.S. has also supported counter-propaganda efforts abroad, financing during the Obama administration up to U.S.$1.3 billion for Europe alone to strengthen resilience against Russian meddling.

VENEZUELA

During his 14-year tenure, Chávez relied on an extensive propaganda campaign to support his government and his ideology, and more recently also made use of social media platforms. He also threatened to silence existing private media organizations if they criticized the government (Guevara, 2018). Following his predecessor, the government of President Nicolás Maduro has intensified a clampdown on the media, blocking the transmission of news channels and websites (Rosati, 2017).

On the other hand, a Pew Research Center study found that Venezuelans primarily use social media to access political news, mobilize protests, and expose government corruption and human rights violations (Guevara, 2018). The opposition to the government has been organizing

itself around social media and has managed to receive attention and support from organizations and politicians also outside the country, receiving support especially from the United States ('Digital Guarimbas', 2017).

Political actors are reported to have used paid online campaigns as a tool to disseminate information. Reports by hackers and data analysts show that campaigns over social media are cheaper and that people can do much more with less money, "sometimes it can take just U.S.$50,000." ('Digital Guarimbas', 2017).

In Venezuela, local organizations report that digital media have increasingly become a source of information and a key factor in determining the future of social movements, specifically how they grow and how powerful they become ('Digital Guarimbas', 2017). Independent digital media, journalists, and citizens actively use digital platforms to access and share critical information (Freedom House, 2013).

Several reports by hackers and data analysts have stated that the opposition has used paid campaigns to promote ongoing anti-government protests, many of which have turned violent. The use of sponsored Twitter hashtags, which they invest money into, has been a recurring strategy, with many of them reaching global trending status ('Digital Guarimbas', 2017). Another growing practice among Venezuelan expats is publicly heckling top-ranking government officials on trips outside the country. The practice, known locally as *escrache*, is filmed on smartphones and is widely shared on services like Instagram, Twitter and WhatsApp. Government officials are aware of this strategy and have made public declarations against it (Rosati, 2017).

The government has invested in surveillance tools, specially tailored to monitor social media. According to Ipys, the director of the Venezuelan Telecommunications Agency (Comisión Nacional de Telecomunicaciones - Conatel), Andrés Eloy Méndez, has announced that the Venezuelan government is acquiring technology to identify, block and censor content, social network profiles, and applications in the country (IPYS, 2017). Also, a presidential decree from May 2017 has mentioned measures to censor and monitor the Internet to prevent "destabilization" campaigns (Freedom House, 2017).

From the government side, Freedom House has reported blocking of political, social, and economic content. Arbitrary website blockings have been reported, notably a handful of sites that provided live coverage of anti-government protests (Freedom House, 2017). Also according to Freedom House, senior officials announced initiatives to regulate the use of social networks,

arguing that they are dangerous and a tool for unconventional warfare (Freedom House, 2017; Méndez, 2017). Private companies have also been hired by government officials to take down negative political content (IPYS, 2017). Freedom House has also reported that:
"According to the investigative journalist Lissette Boon, from *RunRunes*, a company called Eliminalia was tasked with "cleaning up the reputation" of Venezuelan politicians and businesspeople on the web. According to the reporter, in less than 9 months, at least 5 news sites received such requests, under the justification that the right to privacy and reputation were being damaged. Those requests were rejected as a form of censorship." (Freedom House, 2017).

Social media have also been frequently used to spread disinformation, often with the help of bots, to amplify political discourse. This fake news comes from a variety of sources, many of them unidentified, most of them with a vested interest in the current conflict (Latouche, 2017).

They also take many forms, from dubious government claims on official news channels to unverifiable opposition claims that spread online (Viana, 2017).

Government officials often claim that the opposition also uses fake news with the purpose of garnering foreign attention. According to Ernesto Villegas, Venezuela's minister of communications and information: "the efforts of a gigantic media apparatus are put to this narrative of Maduro's government as a massive violator of rights" in an effort to bring about change (Viana, 2017).

The government has allegedly used state-controlled media, "armies of trolls," and encouraged pro-government social media users to harass those with opposing views (Freedom House, 2017; IPYS, 2017). In April 2017, the government announced that it would create "digital militias" by setting up hundreds of points throughout the country to help sign up citizens to social media accounts. According to analysts consulted by Freedom House, the objective of the "militias" would be to increase counter-information and disseminate pro-government messages (Freedom House, 2017). The program, called Gran Movimiento Robinson Digital (Robinson Digital Grand Movement), according to local sources, aims to "produce content, establish strategies and influence social networks." (El Nacional Web, 2017). According to Ipys, the "digital army" of trolls is organized into squads, brigades, and battalions—each squad is formed of one person, who is in charge of 23 social media accounts, and each brigade is formed of 500 people, who handle 11,500 account (IPYS, 2017). According to Freedom House's 2017 report:

> "The NGO Ipys Venezuela shared a May 2017 leak of a presentation from the Ministry for Interior, Justice, and Peace presenting government strategies of military organization and intelligence to

inhibit users from debating on social networks. In March 2017, President Maduro announced the creation of the "Robinson Digital Grand Movement" aiming to "win the war" on social networks by producing content, developing communication strategies, and training people in the use of digital tools".

In Venezuela, authorities have used pro-government Twitter bots to manipulate social media. The Venezuelan President Nicolás Maduro recently became the third-most-retweeted public figure in the world, behind only the king of Saudi Arabia and the pope, mostly thanks to the large number of fake Twitter followers he has (Brooking & Singer, 2016).

In 2017, fake news spread worldwide about a petro cryptocurrency being proposed by President Nicolás Maduro. The Superintendent of Venezuelan Cryptoassets and Related Activities later denied the information (Reuters, 2018). In another widely reported case of fake news, false information circulated claiming that political prisoner Leopoldo López had died on May 3, causing great alarm among the population (Freedom House, 2017).

VIETNAM

In 2013, the Vietnamese government admitted it employed circa 1,000 staff, who engage in online discussions, on social media and forums, and post comments that support the Communist Party's policies. They are referred to as "public opinion shapers." The BBC reported that the head of the Hanoi Propaganda and Education Department, Ho Quang Loi, stated that "Internet polemists" were used to combat "online hostile forces" and that the department managed 400 accounts and 20 microblogs. According to Loi, this digital strategy helped in stopping the spread of negative rumors and blocked opportunities for mass gatherings. The same year, the government introduced a law that banned the discussion of current affairs on the Internet; instead social media and blogs should only be used to share personal information.

In December 2017, Colonel General Nguyen Trong Nghia, who is deputy chairman of the General Political Department of the People's Army, announced that 10,000 "core fighters" are staffed in Force 47, which is responsible for combating false news, "wrongful views" and anti-government content online. Nghia justified this force in light of 62.7% of the population having access to the Internet.

In late 2017, between 54% of the population had active Facebook accounts (52 million). According to *The New York Times,* YouTube and Facebook account for 2/3 of the domestic digital media market. Unlike China which bans foreign social media, the Vietnamese government

allows it and uses it as a platform to disseminate its own media as well as monitor critical content. In early 2017, the information ministry issued a circular to websites, social media sites and apps that have over a million users in Vietnam to work with the authorities to block or remove "toxic" content online. Google partially complied with a request to remove 2,300 videos on YouTube by removing under 1,500. Facebook set up a separate channel to communicate directly with the Communication and Information Ministry to prioritize governmental issues with fake news that circulates as content or as advertisements.

In light of cyber security and fake-news concerns, the government drafted a cyber-security law in June 2017, which is noticeably broad and has been criticized to seek out formal control over social media, requiring foreign technology firms like Google, Facebook, Viber, Uber and Skype, to set up offices and data servers in Vietnam. Nguyen Hong Van of the Vietnam Institute of Information Security argues that domestic data ownership will safeguard the country's cyber security. The law was inspired to "prevent news sites and blogs with bad and dangerous content," according to President Tran Dai Quang, which "undermined the prestige of the leaders of the party and the state." If the firms do not comply, they will not be allowed to offer their services in Vietnam. The draft law received criticism for going beyond cyber security and taking aim at controling content. It will be voted on by the National Assembly in June 2018. On June 8, 2018, the U.S. and Canada urged the Vietnamese government to delay their vote on the Cybersecurity bill so that it can align with international standards. It also concerns activists, whose freedom of expression will be curtailed if the government has access to Vietnamese data on social networking platforms.

ZIMBABWE

The July 2018 presidential election will be the first social media election in Zimbabwe. The last election was in 2013 before social media had taken off. The research network, Afrobarometer, found in 2017 that 84% of Zimbabweans have mobile phones and more than 36% access the Internet on their phones. In June 2017, 850,000 Zimbabweans had accounts on Facebook, some 5% of the population. According to TechZim, over five million citizens use WhatsApp.

The frontrunners are Emmerson (Ed) Mnangagwa of the ruling Zimbabwe African National Union – Patriotic Front (ZANU-PF), who has been in power since the coup d'état in November 2017 and resignation of Robert Mugabe who had been in power for some 37 years; and the youthful Nelson Chamisa of the Movement for Democratic Change (MDC). Mnangagwa wrote an op-ed in *The New York Times*, inviting the international community to invest in Zimbabwe and describing the new Zimbabwe as a "country of hope and opportunity." It has been described as

fake news by critics. The government officially seeks a legitimate and fair election, but critics point out bias in the Zimbabwean Electoral Commission, issues with voter registration and state control of the police and media.

According to Dr. Chipo Dendere, a political scientist at Amherst College and Zimbabwean political commentator for the *Washington Post*, campaign rallies are now virtually closer to citizens at all times thanks to live streaming platforms and people sharing video clips on WhatsApp. According to Newsday, a Zimbabwean news outlet, 60% of registered voters are in the youth category, which will make digital campaigning critical. Campaigners are creating social media arms to reach voters. Dendere comments that one of Mnangagwa's first acts as president was to verify his Facebook and Twitter accounts. While the MDC has traditionally projected itself as the youthful party, ZANU-PF has been ahead in social media, argues Dendere. Much of the social media campaigning is dependent on funding and resources to employ a team dedicated to social media politics.

WhatsApp, Facebook and Twitter are dominant platforms for the election campaigns. WhatsApp has been highlighted as the main platform for circulating news fast. Dendere commented that what has been striking for her as a researcher who is a member in a number of WhatsApp groups, is observing the pace at which she receives messages in various groups with the same message or news story. She states that groups have been, on average, 15 to 50 members and each member is "in anywhere between 5-10 groups," including church and family. This dynamic allows for messages to spread very quickly. While many Zimbabweans use Facebook, to stay up-to-date with election news Dendere points out one has to be in active groups. Twitter for most part is very exclusive for the elite and middle class, who can afford to pay for the bundles.

The ZANU-PF Youth League has been playing an important role in the spreading of #EdhasMyVote on social networking platforms. Cde Pupurai Togarepi, Secretary for Youth Affairs for the ZANU-PF party, commented that the #EDhasMyVote campaign strategy is significant, giving young voters a digital platform to discuss politics on social media. Over 16,500 Facebook users have liked the page on Facebook. ED Pfee (meaning ED, the old one, who is making a comeback) is used and gains popularity, according to Dendere, because it got him in trouble with Robert Mugabe before. The popular hashtag in Chamisa'a campaign is Chamisa Chete Chete (which means 'Only Chamisa'). #Godisinit has also been used, mostly on Twitter; but is less popular that Chamisa Chete Chete. Chinja (the MDC's slogan in Shona) is a popular hashtag for MDC supporters.

The rise of social media challenges the government's stronghold on mainstream and state-run media, which supports the government. For instance, state media has been reported to broadcast rallies of the ruling party, the ZANU-PF, and the president, Emmerson Mnangagwa; while it does not broadcast rallies held by opposition parties. On social media, Dendere states, people tend to call out the state-controlled news agencies, ZBC and *The Herald*. Dendere argues, however, it is unsure what impact the challenging of state-controlled media has in low income and rural areas, where social media has been adopted less and propaganda has the most impact.

Challenging the government online has been careful, following a recent history of arrests and suppression of online activism. The 2017 'Zimbabwe Free the Net' report by Freedom House comments on social media commentary during the 2016 cash crisis. Citizens shared pictures of empty store shelves and price hikes of more than 300% on basic commodities. The *Washington Post* states the hype on social media reminded many Zimbabweans of the 2008–2009 crisis when hyperinflation reached 89.7 sextillion percent. The government responded to growing online activity with a Cyber Security, Threat Detection and Mitigation Ministry in October 2017. Prior, it passed the Cybercrime and Cyber Security Bill 2017. Officially, the bill and ministry seek to tackle cyber bullying and revenge porn, amongst other online activities. However, critics challenge that its motive is to legitimize the surveillance of government critics on social media as well as the increasingly connected populace at large.

A renowned case of digital activism was Pastor Evan Mawarire's #ThisFlag Facebook video in April 2016. The video sparked a campaign which carried through the summer, provoking anti-government protests. During the protests, WhatsApp was made inaccessible for several hours and overnight mobile data prices increased by 500%, according to Freedom House. Mawarire was arrested in September 2016 and faced 20 years in prison for inciting violence against the government, charges which the High Court acquitted him of in November 2017. Mawarire tweeted a selfie in court with the following message: "My fellow citizens it is my absolute pleasure to inform you that I have been acquitted of all charges. Thank you for your prayers and support. Let's join hands in building a better Zimbabwe #ThisFlag." Deprose Muchena, Amnesty International's regional director for Southern Africa, commented in November, "The task for President Mnangagwa now is to ensure that a culture exists in Zimbabwe in which voices from outside his government are free to air their opinions on an equal platform, without fear of facing criminal charges."

SERIES ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of the European Research Council, Computational Propaganda: Investigating the Impact of Algorithms and Bots on Political Discourse in Europe, Proposal 648311, 2015–2020, Philip N. Howard, Principal Investigator. Additional support has been provided by the Ford Foundation. Project activities were approved by the University of Oxford's Research Ethics Committee (CUREC OII C1A15-044). Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funders or the University.

For their assistance and advice on this research, we are grateful to Maud Barrett, Beatriz Kira, Cailean Osborne, Jan Rau and Julia Slupska for collecting the preliminary data and drafting country profiles about social media manipulation for the countries outlined in this report. We are also extremely grateful to Dan Arnaudo, Nicole Au, Crofton Black, Ingrid Brodnig, Chipo Dendre, Monika Glowacki, Dean Jackson, Bence Kollanyi, Andreas Krieg, Vidya Narayana, Lisa-Maria Neudert, Sarah Oh, Anna Pleshakova, Akin Unver, and Syddharth Venkataramakrishnan, as well as the many anonymous experts we consulted for this project. Their country-specific expertise and networks were essential for ensuring the reliability and validity of our data. We thank them for their time and assistance in reviewing country profiles, and for providing us with additional sources, citations and data-points to include in this report.

AUTHOR BIOGRAPHIES

Samantha Bradshaw is a doctoral candidate at the Oxford Internet Institute, University of Oxford, a researcher on the Computational Propaganda project, and a Senior Fellow at the Canadian International Council. Prior to joining the team, she worked at the Centre for International Governance Innovation in Waterloo, Canada, where she was a key member of a small team facilitating the Global Commission on Internet Governance. Samantha's research has been featured in numerous media articles, including the *Washington Post*, *the Financial Times* and *CNN*. Samantha holds an MA in global governance from the Balsillie School of International Affairs, and a joint honors BA in political science and legal studies from the University of Waterloo.

Philip N. Howard is Director of the Oxford Internet Institute, and a statutory Professor at Balliol College, Oxford. He writes about information politics and international affairs, and is the author of eight books, including The Managed Citizen, the Digital Origins of Dictatorship and Democracy, and Pax Technica: How the Internet of Things May Set Us Free or Lock Us Up. He has won multiple "best book" awards, and his research and commentary writing has been featured in the New York Times, Washington Post, and many international media outlets. Foreign Policy magazine named him a "Global Thinker" for 2017 and the National Democratic Institute awarded him their "Democracy Prize" for pioneering the social science of fake news.

UNIVERSITY OF
OXFORD