Table of Contents

Table of Figures

METHODOLOGY NOTES

These are the background case notes complied for The Global Disinformation Disorder: 2019 Inventory of 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 700 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.

ANGOLA

Angola is not considered a free country by Freedom House. Since independence in 1975, the MPLA party has been in power. José Eduardo dos Santos was president for 38 years until João Lourenço took over when he won the election in August 2017. According to monitoring carried out by the African Union, the elections proceeded relatively peacefully and were well organized, however pro-government media, deficiencies in voter registration and the fact that the MPLA used government resources for their campaign gave them an unfair advantage. Opposition parties called the election fraudulent, but the High Court of Angola dismissed the claims and instead accused them of providing fraudulent evidence themselves.

In terms of media and Internet prevalence and access, the Angolan government owns most of the media in the country, and most influential outlets in Angola which are based outside of the country are usually privately owned by MPLA members and function as mouthpieces for the party. Additionally, the MPLA and its members own service providers (TV, radio and increasingly Internet providers) and prevent critical media from reaching Angolan citizens. For example, Zap, which provides TV subscriptions and is owned by Isabel dos Santos, removed a critical Portuguese news channel from their programming from March to June 2017 during campaign season. However, most Angolans are not able to afford TV subscriptions. To date, the Internet remains the least controlled medium, so the most critical voices are found there. However, Internet access remains one of the lowest in the world with a penetration rate of only 13% in 2016 and 45% mobile phone penetration in 2017. Regarding social media, in 2019 Facebook is by far the most popular (about 93% of social media use), followed by Pinterest and YouTube (roughly 2% each), with the platforms Instagram, Twitter and Tumblr making up the last 2% together. Again, the main reason for this is the cost of Internet and mobile phone subscriptions: unlimited access to the Internet costs on average of US$150 a month. For these reasons, conventional media such as print newspapers or the radio are the main sources of information in Angola.

Given that most outlets and service providers are owned by the government or party members, the MPLA has a fair amount of control over what news is broadcast. Furthermore, critical journalists and activists are regularly prosecuted. For example, in March 2018 Rafael Marques de Morias was taken to court for allegedly insulting the country after publishing an article in October 2016 claiming that the then Attorney General João Maria de Souza had engaged in corruption in order to acquire beachfront property. De Morias was acquitted in July 2018 in what has been called a landmark ruling towards more press and speech freedom in Angola. Moreover, the new president Lourenço has replaced several heads of media outlets and urged them to serve the public rather than individual politicians. He is also taking steps against corruption: former governors have been taken to court over money laundering and abuse of power; de Santos family members at the head of several state-owned corporations were fired, and Angolan money held in private accounts outside of the country (reported up to US$30 billion) is mandated for return to the state. However, information on the effect of any steps taken are scarce so it is hard to tell how successful or serious these measures are.

According to the *Freedom on the Net* report, the Angolan government did not censor any online content during the 2017 presidential elections. Nonetheless, the administration has long voiced intentions to do so, as social media bots were moderately influential during campaigning, making up 9% of influencers online. Only journalists and media organizations had more influence on public opinion[1]. 94% of these bots were operated from outside Angola, mainly from South Africa and the United States, and they subsequently moved on to new areas once the vote had been cast. Notably, there is little information available online on the bots involved, so their exact origins and intentions are not well known.

In 2011, the Law on Electronic Communications and Information Company Services was passed. This law allows the government to "intervene when internet service providers jeopardize their social functions or there are situations that gravely compromise the rights of subscriber and users". Technically, the law is supposed to ensure citizens have universal access to information, as well as transparency in public sectors and participatory democracy. However, observers are weary that the law is and has been abused to censor and control service providers. The last known case of online content control was in 2015 when a Facebook user was arrested for criticizing a military general and forced to apologize and remove his post.

Regarding the government's intention to control online content, the Social Communication Legislative Package was launched in January 2017 and was retained by the new president. The package includes a Press Law, Television Law, Broadcast Law and Journalist Code of Conduct to enable the government to control and censor information online. The package includes language stating that all social communication media have the responsibility for ensuring citizens' rights to inform and be informed in accordance with public interest. Critics, including Human Rights Watch, see this as a way to censor online, as it is the state which defines what is in line with 'public interest', and urged the then president dos Santos not to sign the law. In 2015, dos Santos had already announced his intention to pass such laws to better control social media in the interest of Angolan citizens; the first such laws were passed including a regulatory body for social communication (ERCA) in mid-2016. With little legislative oversight, the ERCA can regulate journalists' conduct, investigate online content producers, and suspend or ban websites that do not produce "good journalism". Essentially, this package allows the government to surveil their citizens, although their actual capacity to do so is unknown. At the very least, they are engaging in targeted surveillance of selected individuals. Moreover, the Angolan government is increasingly cooperating with the Chinese on surveillance methods. Finally, in early 2019 it was announced that Angola and Rwanda had signed a deal to cooperate on security and public order in the interest of their citizens. This cooperation reportedly includes the sharing of technical advice and information concerning law enforcement.

[1] Keeping in mind that internet penetration rates are fairly low in the country, so the general reach of the internet (and this its influence) is not as dominant as in more developed countries.

ARGENTINA

According to Freedom House, Argentina has good Internet freedom ranking when compared to other Latin American nations (Shahbaz, 2018), with a score of 28. Nonetheless, the country has reported organized social media attacks against at least 55 activists and journalists – particularly those journalists who reported against the government (for example, outing corruption scandals), and who have been besieged with aggressive tweets. Another report, issued by Reporters Without Borders, points to a significant drop in Argentina's press-freedom ranking. The country dropped from position 52 to 57, since polarization in Argentinean society has led to increased violence towards journalists (Reporters Without Borders, n.d.).

Legislative changes in 2017 included digital advertising as a form of regulated political advertisements; coincidentally or not, the government reduced its budget on advertising by 7%. Journalist Marco Bonelli estimated the investments in 200 million pesos (El Pais, 2019). Amnesty International released a report in which it analyzed tweets engaging with 11 journalists and activists who were attacked intensively on social media. Between October and November 2017, 354,000 tweets engaged with these accounts, and 53.2% of the attacks came from automated accounts (Amnesty International, 2018).

Disinformation in Argentina has been reported on many platforms and strategies vary from 'pro-government' campaigns to discrediting the opposition. Experts report the use of "para-addressee" strategies, where disinformation not only attacked a particular party ("counter-addresse"), but also intended to provide explanatory information to third parties, or "para-addressees" (Aruguete, 2019).

The content of these messages discredited political parties and civil organizations, but often would be directed to public institutions, such as the national healthcare system or the judiciary. One campaign spread the story that chemotherapy was the cause of death of cancer patients, and not cancer itself, promoting conspiracy theories (Slipczuk, 2019).

Disinformation in Argentina peaked in 2018 during the discussion on the legalization of abortion. Argentinian society was strongly polarized between political and civil society groups for and against the legalization of abortion. Fabricated stories reported a legal abortion (abortion is authorized for pregnancies resulting from rape), where the foetus would have "agonized to death on a hospital tray during 10 hours" ("Falso En Las Redes," n.d.). Such disinformation led to the verbal and physical harassment of people involved in the case, including journalists and the judge who authorized the victim's abortion.

Strategies also included promoting a sense of nationalist unity and revolt against a common enemy. One fabricated image made up a quote reportedly by Winston Churchill: "If Argentina ever got organized it would rise and lead Latin America behind it". This story sought not only to instil nationalism, but also to build it upon pre-existing historical tensions with the United Kingdom.

A UK Parliament report has also found that Cambridge Analytica had a local partnership with the SCL group, and that it participated in an anti-Kirchner information campaign.

Alexander Nix, former Cambridge Analytica CEO, confirmed this involvement during a hearing at the UK parliament (House of Commons: Digital, Culture, Media and & Sport Committee, 2019).

Finally, another notable trend was the reference to regional geopolitics in order to address political ideologies. Edited images portrayed former Brazilian president Luis Inacio Lula da Silva watching current president Jair Bolsonaro on TV from a penitentiary ("Falso En Las Redes," n.d.). Fake accounts of Nicolas Maduro were also created on Instagram and Facebook, with an edited image implying the accounts had been verified.

AUSTRALIA

In Australia, social media is a fundamental part of political campaigning. This tendency 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 2018* found that 82% of Australians enlist online sources for their news consumption, while 52% rely on social media. This is also reflected in the growing number of people paying for online news, which rose from 13% in 2017 to 20% in 2018. In an important and quite controversial decision, the Australian government changed its anti-terrorism and security laws in 2018. The decision has been criticized by Human Rights and news media organizations for its far-reaching limits on freedom of expression and other civil liberties and its particular threats to the work of journalists and whistleblowers.

Meanwhile, Australia's government, parties and politicians are increasingly realizing the opportunities and dangers offered by social media. After the 2016 Australian Federal election, dubbed the "Facebook election", the government admitted that they were quite unprepared for the misinformation and foreign trolling activity on social media, such as that from the Russian Internet Research Agency (IRA) during the election (Figure 1). However, Australia is probably not as prominent a target for foreign influencers as, for example, the United States where one of their main goals is to suppress voting, somewhat impossible in Australia because voting is mandatory. Nevertheless, researchers Tom Sears and Mike Jensen have found a continuous cyber troop presence by groups associated with foreign countries, especially the Russian IRA. Additionally, The National Media Research Council highlighted the work of Russian trolls and media outlets like Russia Today and Sputnik which disseminate misinformation. The Joint Committee on Electoral Matters concluded after the 2016 election that cyber manipulation of elections is increasingly becoming an issue both in Australia and internationally. Further, in their report published in November 2018, they recommended the establishment of a permanent taskforce to "prevent and combat cyber manipulation in Australia's democratic process".

In January 2019, the first larger-scale computational campaigning effort was undertaken by United Australia during the Queensland election to the Australian House of Representatives. Financed with at least A$50 million (£27,232,500) from Clive Palmer, who also founded the party, SMS messages were sent to at least 5.4 million Australians (Figures 2 and 3). While many reactions were unfavourable – questioning where Palmer's team got the numbers

from – this activity is technically legal according to the Australian Communication and Media Authority who state on their website: "Australians can be contacted by phone, email, SMS in the lead up to election to seek views and influence your vote". Palmer said that they would continue their campaigning effort including SMS messages and that there was no limit to the amount of money that would be spent.

Even though Australia may not be the most prominent target for foreign influence, there is speculation that the country may face domestic interferences from foreign state actors, mainly China and Russia. In May 2018, the unexpectedly large Twitter following of Senator Kitching was brought to the public's attention after it was revealed that 27% of her followers were most likely fake Russian accounts. It has since been revealed that she is not the only one with large Russian followings. However, she is adamant that her Russian following never exceeded 3%, even after she lost more than 4,500 followers in a Twitter culling of fake accounts in early November 2018. It remains unclear whether she instigated the fake following or a third party did. Similarly, several prominent politicians with fake Twitter following have lost a significant number of followers due to culls carried out by Twitter.

Regarding the influences of China, the International Cyber Policy Institute – part of the Australian Strategic Policy Institute – warned that the 1.5 million monthly users of the Chinese platform WeChat in Australia could fall prey to misinformation and propaganda as the service is controlled by Beijing. To what extent this control encompasses content published outside of China is, however, unclear. WeChat was banned from the Australian Defence Department as a means for communication in 2018 and, in June 2018, an Electoral Integrity Assurance Task Force was set up, involving the Australian Electoral Commission, the Department of Home Affairs, the Australian Cyber Security Centre, and the Australian Security Intelligence Organisation. The task force has worked with Western social media companies to ensure misinformation during elections is minimized, but it has failed to talk to WeChat or its father company, Tencent. On February 18, 2019 news broke that several Australian parties were hacked and, according to the prime minister, the attack originated from a "sophisticated foreign state actor". The Australian Strategic Policy Institute has put China "on the top of their list" of suspects although Russia "would not be ruled out", according to Fergus Hanson who is a cybersecurity expert at the Institute. In March 2019, the Electoral Commissioner expressed his confidence that the attack did not affect the electoral integrity of the country. At the time of writing, no further information or reactions were available as the investigation continues.

Alongside the task force, the Australian government continues to gear up to ensure its cybersecurity standard is of the highest level. The 2018/2019 government budget includes A$9 million (£4,893,363) to be given to the Department of Parliamentary Services over a 4-year period to establish a cybersecurity operations network. Additionally, the Australian Cyber Security Centre has set up a 24/7 cyber newsroom in collaboration with the Crisis Coordination Centre to foster early warning and outreach. The newsroom is supposed to capture all comments and opinions on cybersecurity issues while also engaging with citizens via their Twitter account and with news updates on their website in an effort to "influence the narrative more broadly". Additionally, Australian legislation passed several bills in 2018

to further protect domestic politics from foreign influences. For example, the National Security Legislation Amendment Bill, the Foreign Influence Transparency Scheme Bill and the Telecommunication and other Legislation Amendment Bill. Additionally, Australia is one of the first Western countries to ban Huawei and ZTE from providing technology for the Australian 5G network.

Figure 1: Instagram content by the IRA to elicit engagement and accumulate followers

Source: https://www.abc.net.au/news/science/2019-01-20/instagram-australian-federal-election-russian-misinformation/10717034

Figure 2: SMS sent as part of the United Australia Party's campaign

Source: https://www.dailymail.co.uk/news/article-6600995/Clive-Palmers-campaign-spams-millions-Australian-voters-unsolicited-political-SMS.html

Figure 3: SMS sent as part of the United Australia Party's campaign

Source: https://www.dailymail.co.uk/news/article-6600995/Clive-Palmers-campaign-spams-millions-Australian-voters-unsolicited-political-SMS.html

AUSTRIA

Austria has a comparatively long history of digital media manipulation, especially during election season when party competition becomes the most intense. As early as 2004, the ÖVP called on their campaigners to anonymously publish negative comments on websites. More recently, during both the presidential election of 2016 and the federal election in 2017, several manipulation strategies involving dirty campaigning were used by most of the parties involved. Negative comments about competing parties were disseminated through Facebook and Twitter without identifying the owner's affiliations. The biggest scandal of the 2017 election was probably that of Tal Silberstein, an advisor to the Social Democrats (SPÖ) at the time who created two Facebook pages in October 2016 specifically to mock and undermine (now chancellor) Kurz's campaign. The SPÖ stopped working with Silberstein after he was accused of money laundering in his home country Israel.

Since the 2017 election, a right-wing conservative coalition of the ÖVP and FPÖ has governed Austria. Both parties engaged in dirty campaigning strategies during the election, but it seems neither has faced any serious legal or public repercussions. Rather, both national and international papers are writing about the "phenomenon" of chancellor Sebastian Kurz who "brought the far-right into the mainstream". In early 2019, the social democrat party SPÖ started laying out their campaign strategy for the European Parliament elections in May. In contrast to previous elections, they decided not to repeat dirty campaigning habits of the past and instead proposed a fairness pact to be agreed by all national parties denouncing the use of, for example, personal defamation or social media bots, as well as limiting private donations to campaigns to €10,000 (≈ £8784). Both the ÖVP and FPÖ declined to agree on such a deal, stating that any past scandals relating to donations or dirty campaigning had been solved.

However, recent revelations surrounding the partly state-owned company Post AG suggest that past campaign scandals are not quite as resolved as the government would prefer. Specifically, several newspapers have alleged that Post AG has sold thousands of customer

profiles to several Austrian parties for "geomarketing" purposes. As a form of microtargeting, this information would have allowed parties to personalize their campaign messages based on where people lived as well as information such as their party affiliation. When these allegations became public, Post AG announced they would no longer favour any parties by giving them data and were going to delete all old and archived customer information.

Around the same time, the governing ÖVP party together with their leader (current chancellor) Kurz were engaged in another scandal relating back to the dirty campaigning of 2017 when suspicion arose that Kurz and the ÖVP were exchanging personal data of political opponents and other private citizens with the Austrian Intelligence Services. Both parties have so far not commented on these allegations, however, reports indicate that the information exchanged had been collected over several years and may have significantly influenced the campaign strategy of the ÖVP.

After gaining power in 2017, the current government has made some legislative changes regarding the powers of law enforcement online. In 2018, they passed legislation expanding the state's capacity for online surveillance, allowing law enforcement to infiltrate private communication on, for example, Skype and WhatsApp using state-owned espionage software (Bundestrojaner). Moreover, both opposition parties and several journalists have accused the FPÖ – who have appointed the Home Secretary and whose chairman is now vice-chancellor – of replacing people in key positions at the Federal Office for the Protection of the Constitution and Counterterrorism in an attempt to better control the office and suppress investigations against the extreme right. For example, an article in the outlet *Salzburger Nachrichten* cites a Justice Ministry executive officer who wrote to the Minister of Justice saying that "in some cases, institutions are abused to seek a shift in power in Austria."

These accusations are also reflected in the 2018 *Freedom House Report* which states that Austria's government has frequently been criticized for a lack of transparency, has weak party finance laws, and is failing to adequately regulate lobbying or prevent parliamentary corruption. Regards FPÖ efforts to try to gain control and power in Austria, the Freedom House report also states that, in 2017, Vice Chancellor Stache said that the Austrian Broadcast Corporation (ORF) – partly controlled by the government – needed "optimization" of its objectives. At this point, it is unclear what he meant by this or whether he or the government has taken any steps towards such optimization. Additionally, Austria's media landscape is increasingly dominated by papers with close ties to the FPÖ such as unzensuiert.at and wochenblatt.at. They regularly present FPÖ ministers in a positive light and attack opposition politicians, private citizens, and news media criticizing the FPÖ, sometimes with fictitious stories. Even though these papers more or less publicly state they are not interested in independent journalism but rather want to support right-wing movements in Europe (especially parties such as the FPÖ or German AfD), the great reach of these outlets has nonetheless resulted in several politicians and citizens falling victim to online hate campaigns.

However, evidence continues to mount that the FPÖ – and in particular Strache – have close ties to right-wing movements such as the 'Identitäre Bewegung' (Identity movement) and their leaders, and rely on them for campaign and legislative support. For example, last year several media outlets and social media channels campaigned against the ratification of the UN Migration Pact. Researchers have shown that right-wing organized activity online spiked during the period when country leaders were convening to sign the deal – ultimately Austria was among the countries who did not sign. This is one of the few instances where true computational propaganda was used to steer public opinion, though whether the FPÖ was involved or supported this campaign is uncertain. However, in early 2019, Strache was engaged in a legal scandal when he claimed a picture of him and Patrick Lenart (Figure 4), an important representative of the Identitäre, was fake, denying any direct ties to the movement. Ultimately, Strache had to admit that the picture was authentic. Moreover, postings on Facebook and Twitter show that the FPÖ and the Identitäre have been close for years (Figure 5). For example, members of the movement work within FPÖ offices and support the party at public events.

Figure 4: Picture of Strache and Lenart at a "comfortable get-together"

Source: https://derstandard.at/2000100336603/FPOe-und-Identitaere-Zusammen-auf-Demos-beim-Wirt-und-im

Figure 5: Twitter post exposing the long-lasting connection between Strache and the "Identitären"

Source: https://kontrast.at/strache-identitaere-fussi-foto/

AZERBAIJAN

Azerbaijan's oil wealth has allowed its government to fund large international projects and rebuild its military in recent years (BBC, 2018). This has included extensive investment in social media manipulation (Geybulla and Grigoryeva, 2018). These computational propaganda efforts often centre around political events such as protests, rallies and elections (Geybulla and Muntezir, 2018).

Early 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 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).

IRELI's decline since 2014 is also due to the departure of leading figures in the group (Aslanov and Mardiyev). Following this, various other youth organizations such as the youth branches of the ruling party (Yeni ('New') Azerbaijani Party) have taken over their efforts (Geybulla, 2016; Earle, 2017). Youth groups allow a degree of plausible deniability and their involvement is preferred because they are cheaper and more adept at social media (Earle, 2017). While some trolling efforts are coordinated by youth groups, it is also possible that some young people join in with the harassment independently (DeGeurin, 2018).

In February 2018 President Ilham Aliyev announced a snap presidential election on 11 April 2018, which resulted in Aliyev securing another seven-year term. Given the restrictions on traditional media outlets, social media is used by activists and the opposition to disseminate information and organize campaigns. In the lead-up to the election, pro-government trolls and commentators were prevalent online (Freedom House, 2018). Political trolls often used comments copied and pasted from presidential or government statements – these individuals are said to be ruling party members, civil servants, and other pro-government supporters (Geybulla and Muntezir, 2018). In February 2018, Deputy Prime Minister Ali Ahmadov told members of the party's youth branch to use social media effectively and make necessary "sacrifices" ahead of the election, with many seeing this as an invitation to attack the opposition (Freedom House, 2018). These attacks came from personal social media accounts, but also fake accounts disguised with different names, with the authors of the report noting that, while Azerbaijan is far from Russia's troll factories, it is "catching up" (Geybulla and Muntezir, 2018).

The tactics used by these trolls are unsophisticated. Trolling techniques include spreading rumours, creating cartoons or memes, harassment, and reporting accounts until they are suspended (DeGeurin, 2018). Mass-reporting is a well-documented tactic and was used against opposition media outlet Abzas.net on 27 December 2018, as hundreds of trolls attacked their Facebook page and reported it for violating Facebook's 'community standards' (Geybulla, 2019). Journalist Sevinc Osmanqizi has alleged that the president's Assistant for Public and Political Affairs, Ali Hasanov, is linked to the Internet trolls. She wrote an open letter in April 2017 calling on Hasanov to order "his trolls" to stop attacking Osmanqizi on Facebook and YouTube (Said, 2018). Arzu Geybulla, an Azerbaijani journalist, observed that some trolls are tasked with trolling specific activists or members of civil society (Meisenzahl, 2018).

Azerbaijan's conflict with its neighbour 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). Trolls frequently reference atrocities committed by Armenia or the Armenian conflict to "hijack and distort" conversations regarding Azerbaijan's human rights record (Geybulla, 2016).

Hacking to gain access to social media accounts takes place alongside cyber troopers' use of fake accounts, takedown requests, and blackmail (Geybulla and Kobaidze, 2019). This includes the compromising of social media accounts belonging to dissidents and activists

(Geybulla and Muntezir, 2018). For example, on 24 November 2018, Aziz Karimov – a journalist in Baku – had his Facebook account hacked, resulting in his removal as administrator from several Facebook pages, including Turan News Agency, Azerbaijan's only independent news agency. At the same time, administrators of other Facebook pages were also compromised. For example, Azadliq Radio and Azerbaijan Service for Radio Free Europe lost all their video content (2,000 videos, posts and photos) and 25,000 of its 500,000 followers (Geybulla and Kobaidze, 2019). On 29 January 2018, Meydan TV lost 100,000 subscribers to its various Facebook pages, and all content since 2012 was deleted (Said, 2018). Their website continues to be the subject of regular Distributed Denial of Service (DDoS) attacks. On 19 January 2019, an opposition rally witnessed opposition members being questioned by the police based on the geolocation of their phones which indicated they had attended the rally. As Article 39 of the Law on Communication requires mobile providers to provide government institutions with requested information, many took to social media to blame mobile phone operators for disclosing the names and phone numbers of those customers at the rally.

The authorities are openly discussing ideas on how to prevent users from relying on social media platforms for disseminating news, sharing stories and expressing concerns. They have suggested creating a national social network and closing access to other more popular social media platforms, in order to prevent people from slandering Azerbaijan.

BAHRAIN

Computational propaganda in Bahrain takes place in an environment of existing political repression. In February 2011, the Kingdom of Bahrain witnessed thousands of pro-democracy activists taking to the streets to demand political and social reform. The demonstrations were violently repressed, and repression has continued through criminalizing online criticism of the ruling family and sustained social media attacks (Huffington Post, 2015). Hundreds of websites remain blocked, including *Al-Wasat*, the only independent newspaper in the country, and as of May 2017, Qatari organizations Al Jazeera, Al Sharg, and Al Raya.

Multiple organizations are tasked with monitoring social media. In November 2013, a Cyber Safety Directorate at the Ministry of State for Telecommunication Affairs was launched to monitor websites and social media networks (Freedom House, 2018). Further, the Bahrain Ministry of the Interior stated in July 2018 that, in the interest of safety, security and order, the General Directorate of Anti-corruption and Economic & Electronic Security should monitor "social media accounts that violate the law and harm civil peace and the social fabric" (Ministry of the Interior, 2018).

Bahrain's legal framework bans criticism of the royal family and imposes strict limits on content. The King of Bahrain ratified a law in 2014 that imposed a prison sentence of up to 7 years on anyone who publicly insults him (CNN, 2014); and a man was sentenced to 6 years in prison in August 2017 for retweeting an alleged insult to the king (Freedom House, 2018). Article 70 of the Press Rules and Regulations law deems the publication of false news a crime, and thus any site that is critical of the government is vulnerable to being blocked

by the Telecommunications Regulatory Authority (Freedom House, 2018). From June 2017 to May 2018, 27 people were arrested, detained or prosecuted for their online activities, with seven receiving prison sentences totaling 207 months. Nabeel Rajab, President of the Bahrain Center for Human Rights, was sentenced to 2 years in prison for 'broadcasting fake news' and for tweets alleging torture in Bahraini prisons (BBC, 2017).

There have been reports of Twitter trolls originating in Bahrain. This 'army of trolls' has allegedly been active since 2011, when hundreds of accounts emerged following the February protests. Al Jazeera reporter Gregg Carlstrom tweeted that "Bahrain has by far the hardest-working Twitter trolls of any country I've reported on" (Owen Jones, 2013). There are likely thousands of anonymous accounts with few followers and profile photos signifying support for the regime, such as a photo of the royal family (Owen Jones 2013). Nabeel Rajab, the President of the Bahrain Center for Human Rights, is the recipient of "regular troll attacks on Twitter" which he believes are "from government made accounts" (Huffington Post, 2012). Maryam al-Khawaja, a prominent human rights' defender, reported that "people from the Ministry of the Interior… set up fake accounts" targeting her with abuse by alleging she works for Iran, is a traitor, an extremist, and a liar (Huffington Post, 2012).

Online harassment in the form of doxing (revealing an individual's identity and personal details) is a common tactic in Bahrain. One of the most infamous accounts was 'Hareghum' (@7areghum, created February 2011) an account that disclosed information such as photos of people at anti-government rallies, their addresses, employment information, and contact details. This account was used to report and find information about suspected 'traitors' (Owen Jones, 2013). It allegedly advertised a Ministry of the Interior hotline where individuals could report protesters engaging in anti-government activity directly to the government (Institute for the Future, 2018). Pro-regime supporters have even used Twitter to report suspected 'traitors' to the Ministry of the Interior's official Twitter account (@moi_bahrain).

Stoking sectarian tensions is often the aim of automated bot activity. In June 2016, Marc Owen Jones found that trolls defended the decision to revoke the nationality of the Shiite cleric Isa Qasim ('de-nationalization' is a common tactic in Bahrain), with 50% of tweets in the subsequent period featuring #Bahrain coming from bots (Freedom House, 2018). Owen Jones identified 5,000 sectarian tweets related to this hashtag, originating from 1,800 bot accounts – later suspended by Twitter. The majority of accounts identified were created between February and July 2014 and condemned the 'terrorist' acts by the Shi'a opposition in Iran, using derogatory sectarian terms such as 'rawafid' (meaning 'rejectionists' of the true Islamic faith). A sample of this automated sectarian activity can be seen in Figure 6 (Owen Jones, 2016). The Bahrain Center for Human Rights noted that hundreds of accounts on social networks, particularly Twitter, misrepresented the 2011 protests by calling it sectarian and broadcasting violent videos and deliberately misattributing them to the peaceful February 2011 uprisings. The accounts appear to have originated from the Ministry of the Interior, and a report published by former chancellor of the minister's council Dr Salah Al-Bander had previously documented that the government was funding groups to incite sectarian divisions online (Bahrain Center for Human Rights, 2011).

Bloggers, activists and journalists continue to face harassment and prosecution for what they post online. Some bloggers have been killed, such as Zakariya Rashi Hassan Al Asheri who was tortured to death in prison in 2011, leading others to self-censor and become less critical of the regime (Owen Jones, 2013). Bahrain Watch identified that a number of activists and anonymous online critics were being arrested following malicious links sent from fake Twitter and Facebook accounts that impersonated well-known figures but ultimately revealed the activists' Internet Protocol (IP) address (Bahrain Watch, 2013). It is claimed that the Ministry of the Interior's Cyber Crime Unit was orchestrating these attacks and Bahrain Watch identified more than 120 cases where a government account targeted a Twitter account with the IP-spying link using a public Twitter mention. In the year prior to the report (2012–13), at least 11 people had been imprisoned and charged with insulting the king on Twitter, according to media reports (Bahrain Watch, 2013).

There are multiple media and watchdog reports of the role that Western public relations (PR) firms have had in computational propaganda, reputation management, and surveillance efforts in Bahrain. According to the watchdog Bahrain Watch, the government has hired 18 PR firms for promotional campaigns since February 2011, spending at least US$32 million in contracts (Freedom House, 2018). *The New York Times* reported that two Washington-based firms, Qorvis Communications and Sanitas International, were hired by Bahrain to communicate with Western media and place opinion pieces in major media outlets (NY Times, 2011). The Huffington Post reported that Qorvis Communications, Potomac Square Group, and Bell Pottinger had all been hired to improve the Bahraini government's reputation at home and abroad (Huffington Post, 2017). Qorvis Communications was also discovered to have puppet accounts on Wikipedia, used to alter articles that did not portray their clients in a favourable light, including articles related to Bahrain (Morris, 2013). PR firms reportedly utilized bloggers posing as journalists on pro-government blogs, such as Bahrain Views and Bahrain Independent, as well as faking social media accounts and partisan op-eds (Institute for the Future, 2018). Olton, a UK-based intelligence and PR firm, reportedly received a US$250,000 contract from the Bahrain Economic Development Board as early as 2011 to "develop an electronic system to track international media" (Bahrain Watch). The Index on Censorship reports that Olton possessed software that was able to identify 'ringleaders' through social media, which it notes is particularly concerning as dozens of protesting students were dismissed from university based on evidence that was gathered through their Facebook profiles (Index on Censorship, 2012).

Figure 6: Automated Sectarianism

Source: Owen Jones, 2013

BOSNIA

Bosnia and Herzegovina (henceforth: BiH) is a partly free decentralized parliamentary republic characterized by its fragmented constitutional regime and partisan gridlock among nationalist Bosniak, Serb and Croat communities (Freedom House – Bosnia, 2018). Sensationalist and clickbait content is frequently shared on social media to incite hostility among nationalities and Serbia. For instance, in December 2017, one story by the title "After Bakir defended Bosnia today in front of Vučić and Čović: See what the minister of Serbia has said …" covered fabricated hostilities and confrontations between the Serbian Minister of Defence Aleksander Vulin and Bakir Izetbegović, the BiH politician and member of the tripartite Presidency (CIK Media, 2017).

The stopfake.org website, an initiative to monitor and report fake news in the Ukraine and surrounding countries, reports that most disinformation in BiH is disseminated through local media; however, foreign media, notably the Serbian edition of Sputnik *Sputnik Srbija*, are increasingly active in spreading disinformation through local language radio broadcasts and social media which predominantly disseminates anti-West rhetoric (Stop Fake, 2019).

According to an EU-backed report by the citizens' association Zašto Ne? (Why Not?), foreign influence is most strongly exerted through connections with BiH-based media outlets, which use each other as sources and redistributors of disinformation, forming a disinformation hub used by local and possibly foreign actors to influence public opinion (Cvjetićanin et al, 2019). The hub contains 29 media outlets in total, including 15 of which are located in Serbia and 14 in BiH (12 of the latter in Republika Srpska). According to the Zašto Ne? researchers, more than 60% of online disinformation in BiH is political in nature, targeting the US administration and the EU, whose 'value system' is often portrayed as undesirable for local cultures. Conspiracy theories are the most common form of disinformation, which are mostly focused on individual states, in particular the UK which was accused of conspiring against Republika Srpska and its former president Milorad Dodik (Cvjetićanin et al, 2019).

Disinformation stems from two major actors: 'opportunistic disinformers' operating mostly through anonymous websites and social media accounts with financial gain as the primary motive; and political and state actors who spread disinformation via public and commercial media outlets to mobilize support toward their political agendas (Cvjetićanin et al, 2019). Anonymous websites account for two thirds of disinformation monitored by this research group, spawned by an industrious ecosystem of content production and dissemination on social media. The authors state, "The congruence of media disinformation and specific political interests raises concerns over targeted disinformation campaigns in the online sphere, some related to foreign actors and sources" (Cvjetićanin et al, 2019).

BRAZIL

Brazil started employing online disinformation strategies for political campaigning as early as 2010. Fake accounts were used to propagate pro-Dilma propaganda and false stories about opposition candidate José Serra via blogs and Orkut, Brazil's main social media channel at the time (Gragnani, 2018a). The use of computational propaganda, which has increased year-on-year, was particularly evident in the 2018 elections.

It is important to note that Brazilian authorities were aware of and responsive to the threat of online disinformation, and even introduced standard practices to prevent these strategies from contaminating campaigns. In 2017, the Superior Electoral Court (TSE), the body responsible for running the elections in Brazil, passed an electoral reform that explicitly banned fake accounts, automation, and disinformation for campaign purposes (Tribunal Superior Eleitoral, n.d.).

Nonetheless, fake news was a major concern during the October 2018 presidential elections. Authorities were handling sanctions quickly, and the TSE subpoenaed Facebook to remove over 196 pages that contained false information, including Movimento Brasil Livre's (MBL) page, a political party running for the first time and notorious for their online presence and sharing false information on their portals (Maleronka & Declercq, 2018).

Brazil has shown quite an innovative spirit towards political propaganda and the 2018 election was marked by a few incidences illustrating this characteristic. First, politicians as

social media celebrities and influencers is not a new thing in Brazil. In fact, current President Jair Bolsonaro started doing live-broadcasts and putting videos on YouTube as early as 2016. Interestingly, this approach took on a different direction when certain politicians became carriers of disinformation themselves. The now elected congresswoman Joice Hasselmann spread false information accusing the Workers' Party of having ties with the Hezbollah (Filho & Felizardo, 2018). She also alleged that one of the mainstream media companies had a contract of up to 600 million reais (approximately £130 million) to support the Workers' Party.

In the weeks leading up to the first round of elections, which took place on 3 October 2018, Brazilians were complaining about the large volume of disinformation circulating on social media. However, research by the Oxford Internet Institute showed that traffic on Twitter was particularly low (Machado et al., n.d.). Multiple reports released in the period between the first and second rounds, along with findings from investigative journalists, showed that the bulk of Brazilian disinformation was being disseminated through WhatsApp. Over 120 million Brazilians use the platform regularly, making it an effective way of reaching all of the population (Paulo Higa, 2018). As research has showed, WhatsApp disinformation was at least 11 times greater than junk news on Twitter, and clustered networks of WhatsApp groups were used to disseminate disinformation throughout thousands of users (Machado, Kira, Narayanan, Kollanyi, & Howard, 2019).

Further research also showed that the Brazilian public was consuming disinformation not only in the form of news articles, but also via audiovisual content, such as hoaxes, false audio testimonies and even edited videos. One piece of research analyzed the top 50 images being disseminated on 347 WhatsApp groups, where a staggering 107,256 images were shared (Chico Marés e Clara Becker, 2018). Only four of the analyzed images were in fact genuine. YouTube and Facebook links were being shared to disseminate fake news over WhatsApp.

As investigative journalist Patricia Mello later revealed, Brazilian firms had contracts with advertising agencies in the United States to promote digital advertising in Brazil (Campos Mello, 2018). Contracts with one of the major supporters, the company Havan, was up to 12 million reais (approximately US$3 million). More research revealed that data brokerage was being used to promote targeted ads in the campaign (Bruna Martins dos Santos & Joana Varon, 2018). Credit companies such as Serasa Experian sold databases for targeted advertisement, and even leaked databases from phone companies were also used by parties to target users over WhatsApp. At least four companies – Yacows, Quickmovile, SMS Market, and Croc Services – were involved in digital advertising for political parties. Sending out WhatsApp messages cost from 0.08 to 0.12 reais for databases owned by the campaigning party and up to 0.40 reais for databases that belong to an advertising companies (Campos Mello, 2018).

Despite peaking during the 2018 Brazilian elections, disinformation strategies did not die out after the campaign. As Jair Bolsonaro has assumed presidency, he has repeatedly used his Twitter account to disseminate propaganda, attack journalists, and even to divert public

debate. Famously, the president released explicit footage revealing "the truth" about Carnival during the 2019 Carnival festivities (Ernesto Londoño, 2019).

Disinformation has also played a major role outside the campaign, including during the 2018 truckers' strike, which paralyzed truck transportation throughout the country and caused billions of dollars' worth of damages to the country. Another tragic case was a rumour that alleged that former councilwoman Marielle Franco was the wife of one of Rio de Janeiro's druglords. This rumour was released 2 hours after her homicide in March 2018 (Gragnani, 2018b). This disinformation has been released mostly over WhatsApp, and the company has reacted to pressure from authorities and civil society to reduce the virality of their messaging system. Nonetheless, some actors seem to be inviting users to other platforms such as Gab and Telegram, where there is no control over disinformation or even dialogue with the platform holders (DFRLab, 2018).

CHINA

The disinformation landscape in China is particularly complex. A wide-range of state and state-sponsored actors actively use computational propaganda as a tool of censorship and control. In 2019, the Chinese government launched an extensive disinformation campaign against the democracy protestors in Hong Kong. Using Facebook, Twitter and YouTube—platforms that are blocked in China—government officials used social media to spread disinformation, sow discord, and undermine the legitimacy of the activists (Stewart, 2019). Twitter took down more than 900 accounts associated with Chinese authorities (Twitter Safety, 2019), while Facebook identified seven pages, three groups and five Facebook accounts involved in "coordinated inauthentic behaviour" related to the protests in Hong Kong (Facebook Newsroom, 2019). YouTube also removed 210 channels that were actively spreading disinformation about the protest (YouTube 2019). All of these accounts were coordinating to portray the Hong Kong protestors as violent extremists in order to undermine their support internationally.

There have also been reports of Chinese interference in elections in Taiwan and in the United States. 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). Similarly, digital footprints of Chinese propaganda campaigns have been noted in Taiwan during elections (Horton 2018), but also around particular issues such as reunification (Bolsover 2017, Spencer 2019).

In contrast to information operations abroad, China employs an extensive censorship regime domestically. The Chinese government often uses the term "fake news" to delegitimise criticism of the state. The discourse of 'fake news' is used to crack down on dissident voices or discredit opinions that confront the government. 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 rumours 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).
The weaponization of “fake news” has been used in conjunction with more crude tools of censorship, and there is a large body of evidence demonstrating how the Chinese government has actively blocked critical content for many years (Deibert 2013, Wright 2013). These filtering technologies block major global news sources, like the New York Times, as well as social media platforms like Twitter, Instagram and Facebook (Freedom House, 2018). Other kinds of content—such as conversations generated by two experimental “chatbots” on Tencent QQ—have also been censored by the Chinese government for voicing criticism of the government (Allen, 2017). During the 2019 Hong Kong Protests, while manipulated content was pushed through Western social media channels like Twitter, Facebook and YouTube, Chinese authorities simultaneously censored any content and state-owned platforms, including the global video-sharing platform, TikTok (Coleman 2019).

Computational propaganda and the manipulation of online content is often used alongside traditional forms of censorship and information control. As early as 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). Although the existence of the “50-Cent Army” has been debunked, there is a growing body of empirical evidence that large-scale manipulation of social media comments does exist within China.

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” (Sonnad, 2014). 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 video, and various internet forums, including working links to the actual posts (Sonnad, 2014). According to a research published by King, Pan and Roberts (2017) rather than debating critics directly, the Chinese government tries to derail conversation on social media it views as dangerous.

King, Pan and Roberts (2017) estimate that the Chinese government fabricates and posts about 448 million social media comments a year. The researchers found that the Chinese regime’s strategy is to avoid arguing with sceptics 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 doesn’t 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).

In addition to government actors involved in the manipulation of social media, there is evidence to suggest that private companies 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 literally called "wǎngluò shuǐjūn," or Network Navy (網絡水軍—literally, "network water army"). These loose organizations of hundreds or thousands of people recruited through sites targeted at "leisure workers", similar to Mechanical Turk jobs, have been offering services online.

Network Navies offer a variety of services including boosting clients' websites on search engines for specific keywords along with general brand promotion and marketing (Gallagher, 2018). They 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, and offer networks of fake accounts to amplify messages on social media services such as WeChat, the Weibo micro-blogging 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 centre 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).

Working as part of the Network Navy can generate income for users who participate. According to Ars Technica, one website operator said he made about 4,000 yuan ($636 US) 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 $640,000 US) 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).

Another private actor operating in China is the American-based company Devumi, which sells Twitter followers and retweets to celebrities, businesses and anyone who wants to appear more popular or exert influence online. 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). 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).

Citizens and youth groups also contribute to computational propaganda campaigns in China. 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 organisation of a nationalist volunteer social media army, known as "little pink", or xiao fenhong, a name derived from the colour 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 co-ordinate "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 organised 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).

Many members of the "little pink" army belong to the "Emperor's Board", an online forum followed by 29m people, where "crusades" are co-ordinated. China's troll army also organises 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).

COLOMBIA

Colombia scores 31 in the Freedom House's *Freedom on the Net 2018* report and is considered to be a partly free country. Disinformation strategies are not a new concern in Colombia. Since 2016, when a referendum was held regarding the proposed peace deal between the government and the Revolutionary Armed Forces of Colombia (FARC), disinformation on WhatsApp was already active (Pablo Medina Uribe, 2018). Some academics attribute the narrow victory of 0.2% of the 'No' vote to disinformation, since polls had predicted 'Yes' votes to win by a large margin (Argüello, n.d.). However, the then president Juan Manuel Santos amended the peace treaty with the FARC and it was approved by Congress.

In the same year, a hacker named Andrés Sepúlveda was arrested for illegally accessing government information (Watts, 2019). He confessed to his crimes, which included hacking and promoting disinformation campaigns in many Latin American countries. The practices he reported included illegally accessing confidential documents, trolling, and even coordinating over 30,000 fake accounts on Twitter to promote disinformation. He has also

said he had been hired by government parties to promote smear campaigns against the opposition and to generate discredit.

These reports indicated that there had been organized cyber troop activity in Latin America for many years, especially involving politicians, government, and parties. Such activity ranged across all main platforms, including YouTube, WhatsApp, Twitter and Facebook. In fact, the penetration of these platforms is so high that 87.3% of Internet users in Colombia are on WhatsApp and 87.5% are on Facebook (Gobierno de Colombia, n.d.).

Disinformation during the 2018 elections followed this organizational pattern and raised concerns of authorities and competing parties. A Twitter user identified a network of websites that was used, until early 2018, to disseminate political propaganda. This comprised a chain of websites of varying themes, including pets, cars, beauty, and other topics (Serrano, 2018). As the electoral campaign drew nearer, the websites introduced political agendas into their content production. For example, one article on a pet website promoted a candidate's agenda on global warming issues. Similar propaganda articles were released throughout the entire network of websites. It transpired that the chain of sites was detained by an advertisement company called Emotions Media Group, which was hired to promote content production and digital impulsion on behalf of a given party.

These rampant assaults of disinformation led the candidates to sign a "fake news non-aggression treaty" (Politica El Tiempo, 2018), where the candidates agreed to respect each other during campaigns and repudiate any form of violence of disinformation against their adversaries. Nonetheless, misinformation continued to rage through social media. In June 2018, before the second round of elections, new hoaxes circulated online. One hoax (with 23,000 shares) portrayed former adult actress Mia Khalifa as candidate Petro's illegitimate daughter, supposedly voting against her father. Another featured candidate Gustavo Petro sharing a hoax which portrayed another adult actor as the apparent winner of a quantum physics prize. While it is unclear if either hoax had political intentions, they obtained thousands of shares and followers – the Khalifa post registered 23,000 shares (Penarredonda, 2018). The spread of misinformation remains very high.

It is worth noting that with the FARC peace treaty and the transformation of the paramilitary group into a political party, social media has become the main turf for disputes in Colombia (Worley, 2018). As reported, FARC replaced "AK's for Tweets" (AK-47 being their preferred rifle) and started campaigns criticizing the establishment and promoting its image (Iriarte, 2017). Conversely, other political parties use the Internet to attack the activities of former FARC politicians and stain their personal images.
One trending disinformation strategy in Latin American is to cast doubt on the trustworthiness of the electoral process. These conspiracy theories were repeated in Brazil, Mexico, and in Colombia, but in no case has there been any evidence of large-scale tampering with ballots. Though fake accounts of politicians and celebrities have been identified, it seems that the bulk of disinformation in Colombia has been disseminated organically (Argüello, n.d.). Many politicians have engaged with fake content and even shared websites that spread blatantly polarizing and false information, such as voces.com.co and oiganoticias.com.

Fact-checkers' efforts were in vain during the elections because the volume of disinformation was way beyond what these bodies could handle. They were often flooded with requests from users and it was very difficult to select what to verify.

CROATIA

Croatia is a free democracy with a 'partly free' press and polarized media landscape (Freedom House – Croatia, 2018). In March 2019, thousands of citizens protested in support of media freedom after the Croatian Journalist Association (HND) exposed 1,160 lawsuits initiated by public figures against journalists, which was seen as harassment and infringement on journalistic freedom (Peruško, 2019).

According to the 2019 Reuters *Digital News Report*, 89% of Internet users in Croatia access their news online via computers (68%) and smartphones (76%). Social media networks are increasingly popular for receiving news, including Facebook (56%), YouTube (28%), WhatsApp (14%) and Viber (13%). However, Croatians have very little trust (30%) in the veracity of news on social media (Peruško, 2019), and disinformation is key in this lack of trust. According to a Eurobarometer survey in March 2018, 47% of Croatians encounter fake news every day and 29% encounter fake news at least once a week (Vejković, 2019). As part of the wider European action plan against disinformation, initiated by the European Commission, the government plans to create a national contact point and real-time alert system that will work with other EU member states to counter online disinformation (Veljković, 2019).

Disinformation is spread on social media networks, in particular Facebook, often with the European Union and migrants as targets. For instance, at an event in Berlin in November 2018, the news site dnevno.hr reported the false story that Merkel had said "member states today must be ready to surrender their sovereignty". This story came from a known 'fake news' website called NewsPunch, which regularly publishes anti-EU and anti-immigration stories. Campaigns have also targeted refugees and migrants. In 2018, Faktograf, a fact-checking site in Croatia, reported that a politician from the Eurosceptic populist party Zivi Zid was systematically spreading disinformation about migrants on Facebook. Online news sites, such as 4dportal and Conflict, have distributed false stories about migrant-led crime, violence and rape via their Facebook pages, including recycling in 2018 a 2015 story about how "migrants tried to smuggle as many as 52 tons of weapons and ammunition across the border! What's on our way?" (DZ, 2018). Given these stories, the online news site warns of fake news that seeks to manipulate readers by eliciting fear (DZ, 2018). Trolls are active on Facebook, Instagram and Twitter who comment, share and like disinformation content, and troll journalists, politicians and civil society groups. Trolls on Facebook are characterized by having no or few friends, photos and posts, but demonstrate active participation in commentary on news.

CUBA

As a one-party communist state, Cuba has no political pluralism, suppresses dissent, and severely restricts freedom of the press, assembly, speech and association. The government

controls virtually all media outlets in Cuba and restricts access to outside information. A small group of independent media outlets in the country are deemed illegal and considered "enemy propaganda". Access to the Internet is very limited, most citizens can only access the government-controlled national Intranet. Wi-Fi spots are gradually being established, as well as home access to the Internet, but again this is done by the state telecommunications company and does not allow for international Internet access. Moreover, the prices for this access are likely too high for most Cubans: the average wage in Cuba is about US$25 monthly and costs are probably going to be at least US$15 for 30 hours of Internet access.

Dissent and critique (both on- and offline) is suppressed and punished by the state. Cuba remains one of the most unconnected and repressive countries with regards to communication and information technologies: a recent study by the Open Observatory of Network Interference (OONI) found 41 blocked sites on the island's Internet, while foreign Internet services remain virtually inaccessible. Nevertheless, Cuba differs from other repressive regimes as its main strategy to keep citizens away from unwanted content is to make the necessary technology unavailable, rather than employing sophisticated blocking techniques. They do, however, have a fairly well-developed system to filter domestic SMS containing words such as "democracy", "dictatorship", or "human rights". Additionally, the Cuban government tries to control the online public narrative by launching copy-cat versions of worldwide services such as Wikipedia, Facebook and Twitter. This way, citizens are only exposed to highly curated versions of each page: in 2010 the Cuban Wikipedia Ecured was launched, in 2013 a Cuban Facebook La Tendedera followed, and finally in 2015 a blogging page knows as Reflejos was launched.

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 by resorting to the techniques of modern black propaganda: falsifying statements, editing video and audio tapes, and making photo montages that are then disseminated via 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 rumours, attacking the credibility of those who they find "uncomfortable," and sowing disinformation lines that justify the repressive operations of their institution (FHRC, 2017).

Meanwhile, evidence is also suggesting 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).

Finally, interference from foreign governments in domestic political debate has been recorded for some time. In 2014, news outlets reported that the US government had developed and implemented an app aimed at undermining the Cuban government. According to the news articles, the US 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.". Documents show that the US 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.". Additionally, reports (e.g. The Real News Network, *Miami New Times*) accuse the US Office of Cuba Broadcasting of using "native" and "non-branded" accounts on Facebook and YouTube to spread right-wing, pro-US, pro-capitalist propaganda in Cuba. A spokesperson of the Broadcasting Board of Governors said the project never took off, though this statement appears unverifiable at the moment.

CZECH REPUBLIC

The Czech Republic is a 'free' democracy with a competitive media ecosystem, where concerns about media stem more from the influence of private business than government (Freedom House Report 2018). Significant concerns surround the influence of Andrej Babiš, billionaire prime minister and leader of the Movement of Dissatisfied Citizens (ANO) party, whose media holdings are held in a trust led by a close associate. In 2017, Babiš was exposed – by leaked recordings – for interfering with the editorial policy of the daily newspaper *MF Dnes* by instructing the publication of articles damaging to political rivals (Freedom House Report 2018, 2018). Amid concerns about the political independence of traditional media, trust levels in news are low at 33% (Štětka, 2019). Increasingly Czechs get news online, through computers (71%) and smartphones (51%) and, in particular, on social media networks, including Facebook (50%), YouTube (26%) and Facebook Messenger (17%) (Štětka, 2019). Trust in news on social media networks is even lower at 20%, according to the 2019 Reuters *Digital News Survey* (Štětka, 2019).

Ahead of the European Elections in May 2019, the core misinformation sources in the Czech Republic were identified as coming from Russia but also increasingly China, which spreads content seeking to improve public opinion about long-term diplomacy and trade with China (České Noviny, 2019). According to František Vrabel, fake news analyst and consultant to

NATO, the Czech president Miloš Zeman is an active promoter of Russian disinformation campaigns who engages with and shares fake news stories, for example that the EU limits the sovereignty of its members or that the US is a global gendarme that begins wars where it can (České Noviny, 2019). According to one study, the cumulative impact in the Czech Republic of pro-Russian disinformation via online news sites has been significant, generating roughly half the page views received by the third most-visited news site, News List (Urban, 2019). The most-visited 'disinformation site' is Parlamentní listy with 8.5 million views per month, which is led by Senator Ivo Valenta. The second most-visited site is a Czech version of Sputnik, followed by AC24.cz and Aeronet.cz which receive around one million views per month. These websites are considered the 'big four' disinformation sites, which the study claims make up 85% of domestic Russian disinformation websites (Urban, 2019). Before the European Election, the director of the Czech national intelligence service, Michal Koudelka, warned the biggest threats of disinformation come from Russia and that increasingly strategies had been transforming, from visibly foreign-led campaigns to more discrete, domestic operations (Kalenský, 2019). The core targets of disinformation campaigns are the European Union – described as "the tool of totalitarianism and controlled globalization" (Countercurrent), "parasitic over-European structure" (Curious), and "a bureaucratic totalitarian apparatus" (Free newspaper) – and the United States (Kalenský, 2019).

ECUADOR

Before the elections in Ecuador in 2017, there was substantial evidence that former President Correa had established a series of troll farms in order to spread pro-government messages, discredit opposition, and supress political dissent and journalistic freedom. The strategies, tools, capacity and organization form of these activities are highlighted in the 2017 and 2018 cyber troops report in greater detail.

Beyond government run troll-farms, the 2017 elections in Ecuador also experienced evidence of social media manipulation. Disinformation was used as a campaign tool by both major parties to support their position as well as undermine the opposition. One prominent example, highlighted by Snopes, described rumours about Correa's party tampering with votes via messages spread on the popular messaging platform, WhatsApp.

Following the defeat of former President Correa to newly elected President Lenin Moreno, there have been few instances of disinformation, and even fewer instances of coordinated campaigns on social media. Although there is no empirical evidence of disinformation or troll farms currently operating in Ecuador, there are rumours that they exist. However, according to Snopes, this could be a signal of how "confidence in information has suffered since Correa's time in office."

EGYPT

Over the past year, the Egyptian government has increased its repressive hold over freedom of information on the Internet. It has extended existing policies of censorship and

surveillance while displaying evidence of limited and relatively unsophisticated computational propaganda techniques. Following the overthrow of President Mohamed Morsi in 2013, President Abdel Fattah el-Sisi has increasingly utilized social media and Internet controls in pursuit of political stability.

Fake social media accounts remain a problem in Egypt. A *Reporters Without Borders* report on online harassment (2018) notes that many Egyptian journalists' social media accounts are increasingly being shut down by "the regime's online armies". The Twitter account of BBC Cairo correspondent Waël Hussein was blocked, and fake accounts began to disseminate content under his name (Reporters Without Borders, 2018). Fake accounts are often more popular than legitimate accounts – according to *The Arab Weekly*, Education Minister Mahmoud Abo el-Nasr's fake account had 80,000 followers compared to the 55,000 followers on his official page. They further reported that "hardly a week passes by" without a press release from the Egyptian Cabinet's Information and Decision Support Centre denying news or information released through fake government accounts (The Arab Weekly, 2019).

Disinformation has proliferated on social media. Following the removal of President Morsi in 2013, there was a surge of disinformation on Facebook and Twitter as both opponents and supporters of the ousted president spread rumours, fabricated images, and created fake accounts. For example, the Facebook page of Egypt's Freedom and Justice Party (FJP), the political arm of the Muslim Brotherhood, posted old photos of children killed in Syria, blaming the Egyptian Army and claiming the photos were from Egyptian protests (Al Arabiya, 2013; BBC, 2013). The volume of disinformation has led to verification pages, such as 'Da Begad?' or 'Is This Real?' to fact-check by verifying posts, images, and videos (BBC, 2013).

There have been limited reports of coordinated social media activity for content removal and trolling. In March 2018, the Facebook page of opposition TV network Watan and a Muslim Brotherhood affiliated page were removed – reported to be a result of government supporters reporting the pages to Facebook for violating their terms of service (Freedom House, 2018). Egyptian authorities are said to have organized "troll armies" that deployed abusive language, threats, harassed people online, and bullied critics – particularly women (WIRED, 2018). Aya Nader, who reports on human rights issues for Al-Monitor and Open Democracy, stated that she has to "think twice" before writing a story or conducting an interview, and has considered writing under an alias – "the online electronic armies or trolls have a great role in that, I have been named and shamed [for writing content that's critical of the government]" (Committee to Protect Journalists, 2017).

Coordinated fake accounts appear to amplify pro- and anti-regime political content. The BBC discovered that while posts by the official Twitter account of President Abdel Fattah al-Sisi attracted an average of 2,000 to 3,000 likes each, many of these accounts appeared suspicious. For example, their activities appeared only to promote pro-Sisi posts, suggesting coordinated activity to make posts more visible (BBC, 2018). In June and July 2018, hashtags such as #El-Sisi_Zaemy_Waftakher (El-Sisi is my leader and I'm proud), as well as opposition hashtags such as #Erhal_Yasisi (Sisi, leave) were posted thousands of times.

Both hashtags contained evidence of organic and inorganic activity, with accounts supporting 'Sisi_leave' coming from those that also tweet about Palestine and are in support of the Muslim Brotherhood (Karan, 2018).

Computational propaganda efforts must be viewed in the context of existing Internet controls and censorship efforts to maintain political stability. While in 2015 only two websites were blocked, around 500 websites were blocked as of July 2018, including many independent media and human rights' groups. Hundreds of proxy tools and Virtual Private Networks that circumvent state censorship have also been criminalized by law (Freedom House, 2018). The government has centralized the Internet infrastructure and fibre-optic cables, creating highly controllable choke points (Freedom House, 2017). Access to particular social media networks are limited during periods of instability, such as during the Arab Spring in 2011. 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" (Freedom House, 2017).

Online censorship increased in the second half of 2018. President Abdel Fattah al-Sisi won a second term in the March 2018 presidential election and introduced pieces of legislation related to the Internet. The most controversial law, enacted in July 2018, focused on the spreading of 'false news' online. The result of the new 'false news' legislation is that any social media accounts or blogs, such as on Facebook and Twitter, with more than 5,000 followers will be treated as media outlets. This makes them liable for publishing 'false news,' which remains undefined and subject to interpretation by the newly formed Supreme Council for the Administration of the Media. In July 2018, President Sisi said in a public speech that fake news had become so prolific that in just 3 months the government had "identified 21,000 rumours" that sought to undermine the government. In September, the government commenced with arresting people under 'false news' charges; a human rights' activist, Amal Fathy, was sentenced to 2 years in prison on charges of "spreading false news", a result of posting a video on Facebook which criticized the government for the country's levels of sexual harassment (Guardian, 2018). Critics argue this loosely defined law will be used to increase repression. In August 2018, the president also signed the Law on Combating Cybercrimes, which created a legal framework to block websites deemed a threat to national security or the economy, as well as criminalizing VPNs (Freedom House, 2018). Individuals who visit banned websites may be jailed for up to one year, and ISPs (Internet Service Providers) are required to hold browsing data and disclose it to security forces upon request (Freedom House, 2018). These efforts have further been supported by the creation of the government's Media and Rumour Monitoring Unit, headed by Naaym Saad Zaghloul, and the creation of a hotline in March 2018 for citizens to report fake news.

Online harassment and censorship is targeted strongly at journalists. Egypt is ranked 161 out of 180 nations in the latest Press Freedom Index, and many Egyptian journalists have registered their opposition to the new laws (Lowy Institute, 2018). The Committee to Protect Journalists, in their annual census of globally imprisoned journalists, found that of the 28 journalists jailed on 'fake news' charges globally, 19 of these were in Egypt (LA Times, 2018). Reporters Without Borders went as far as saying that Egypt is now "one of the

world's biggest prisons for journalists". An Egyptian economist and author, Abdel-Khaleq Farouq, was arrested in October 2018 for his book titled *Is Egypt Really a Poor County?* for publishing 'fake news' that challenged President el-Sisi's economic policies. While the detention of journalists is not new, the Committee to Protect Journalists has called the influx of detentions "fresh waves of repression", particularly under the new justification of 'false news' charges.

ETHIOPIA

The Ethiopian Government maintains a repressive regime over its political space, freedom of expression and marginalized ethnic groups. Corruption, protest crackdowns and human rights abuses mark Ethiopia's recent political history. In May 2015, the ruling party, the Ethiopian People's Revolutionary Democratic Front, won 100% of parliamentary seats; the same year, the government cracked down on opposition political parties, journalists and peaceful protestors, the Human Rights Watch reports. Following the elections beginning in 2015 and carrying through 2016, protests took place in the Oromia region (home to the largest ethnic group, the Oroma, who make up circa 40% of the population) and the Amhara region. State forces used fatal methods to contain demonstrations, killing hundreds of protestors.

In October 2016, the Government declared a state of emergency following the destruction of government buildings and private property by protesting youths. The state of emergency introduced further restrictions on freedom of expression, association and peaceful assembly while legitimizing the government's violent tactics of (information) control. The state of emergency lasted 10 months from October 2016 to August 2017, a period marked by mass arrests and restrictions on independent media and social media.The Government shut down mobile networks for almost two months, thereby controlling spread of information digitally, and removed space for critical voices through punishing critical opinions, which was exacerbated by the conduct of arrests without court orders. According to government figures, over 21,000 citizens were detained and sent to "rehabilitated camps".

A recent report by the Human Rights Watch details that the state of emergency increased government controls over the media landscape: information independent of state-run media became increasingly difficult to access and journalists face a choice of self-censorship, arrest or exile. Reportedly, at least 85 journalists have sought exile since 2010 and a number remain in prison for breaking the anti-terrorism law. Human Rights Watch goes further reporting diaspora journalism, in particular television, has been targeted; for instance, the Ethiopian Satellite Television and the Oromia Media Network stations have been banned under the anti-terrorism law.

According to the most recent Freedom House report, the Internet penetration rate is 15.4% of a population of over 102 million, which makes it one of the least connected countries in the world. The telecommunications infrastructure is undeveloped and high prices for Internet access are controlled by state-owned Ethio Telecom which holds a monopoly over SIM cards, which citizens must register for, and since the 2016 protests they have

announced plans to require mobile phones to be purchased from Ethiopian companies complemented by a tracking system for all mobiles.

The government restricts access to social media and websites containing information deemed subversive. In December 2017, amid anti-government protests, the Government blocked access to Facebook, Twitter and YouTube among other social media. WhatsApp and Telegram, however, were not blocked. According to Moses Karanja of Citizen Lab, network scans of Ethio Telecom confirm that the websites were made inaccessible. Internet shutdowns take place frequently; for instance, in June 2017, during the national exams period, the Internet was shut down across the nation and between May 30 and June 8, all telecom networks were shut down following the conviction of two human rights activists for online expression in May 2017 (Freedom House report). In June 2016, the *Computer Crime Proclamation* was passed, criminalizing amongst others digital content that "incites fear, violence, chaos or conflict among people" and legitimizing interception of digital communications. Freedom House reports on a lack of transparency about how such control is operationalized, exacerbated by the fact that the government denies it censors the Internet.

The 2017 Freedom House report lists the following among a number of persecution cases for online activities:

- October 2016: Seyoum Teshome, academic and blogger, published an article about the Oromia protest in *The New York Times;* Teshome was imprisoned and tortured for 3 months
- November 2016: Anania Sorri, Daniel Shibeshi and Elias Gebru, activists and a journalist, were arrested for positing images on social media gesturing support for the protest movement. Protest gestures were banned during the state of emergency
- December 2016: 7 musicians were arrested and held without charge until June 2017 for their involvement in a YouTube music video. In June 2017, they were charged with terrorism for inciting protests
- May 2017: Yonatan Tesfaye, a prominent activist, was declared guilty of terrorism based on Facebook posts critical of how the government responded to the Oromia protests. Tesfaye was sentenced to 6.5 years in prison. Tesfaye's Twitter handle has been active since, provoking suspicions that officials are using his account to monitor other dissidents or encourage them to break the law.
- May 2017: Getachew Shiferaw, editor-in-chief of *Negere Ethiopia,* the opposition news outlet, was sentenced to 1.5 years in prison on subversion charges for Facebook comments that "endorse[d]" an exiled journalist. He has been released since.

Critics of the government who have sought exile abroad have been targeted for producing critical content about the government. According to WIRED Magazine, dissidents in 20 countries, including Germany, USA and Canada, were targeted through spyware embedded in emails containing a malicious link. The PC Surveillance System spyware was produced by Cyberbit, an Israeli firm and subsidiary of the defence contractor Elbit Systems.

GEORGIA

Georgia is a 'partly free' democracy with a competitive but frequently partisan media landscape which, according to civil society groups, is increasingly at risk from the political influence of oligarchs and leaders of the governing Georgian Dream party (Freedom House – Georgia, 2019). In 2017, civil society groups released a joint statement expressing concern about the recruitment of political allies of Bidzina Ivanishvili, chairman of the Georgian Dream party, to senior positions of the Georgian Public Broadcaster, which the statement describes as decreasingly critical of the government (Freedom House – Georgia, 2019). In 2018, threats to the Georgian media ecosystem were brought to international attention when a long-standing legal dispute about the ownership of Rustavi 2, an opposition-aligned TV station, involved the European Court of Human Rights (ECHR) (Freedom House – Georgia, 2019). In 2019, the ECHR ruled against the owners of the TV channel in a controversial ruling (Antidze, 2019). According to the Media Development Fund (MDF) – a Georgian NGO – in addition to concerns about threats to the diversity of traditional media in light of government influence, the government frequently employs troll factories to mobilize public opinion on social media, and to criticize entities ranging from TBC bank to NGOs to media outlets. Social media trolls and sponsored posts are reportedly particularly prevalent during anti-government protests, where content is spread to share pro-government information (MDF, 2019).

The former prime minister Kvirkashvili, who resigned in the summer of 2018 due to disagreements with the party chairman Ivanishvili, was exposed for buying likes to promote his posts on Facebook. In June 2018, Kvirkashvili and the government more broadly were found to have paid for thousands of likes from accounts in India, Bangladesh, Vietnam and Pakistan, among others, after thousands of users commented "haha" on a Facebook post criticizing NGOs (Gvadzabia, 2018). According to local news sources, the post rejected an ultimatum set by NGOs for the Justice Minister, Tea Tsulukiani, to resign for failing to nominate a candidate for the post of Chief Prosecutor. Kvirkashvili wrote that NGOs should not go beyond the mandate of their activities, and criticized NGOs for lacking in transparency themselves. The post spawned negative responses in comments and likes, which were countered by the allegedly bought likes (Gvadzabia, 2018). According to the MDF, pro-government trolls on Facebook frequently target NGOs and journalists. Between March 4 and April 5, 2019, 15 pro-government trolls were observed. Eleven users matched stolen identities from Russian social networks, Odnoklassniki and VKontake; two users were accounts named after characters in a TV series; one was a Facebook user; and one was a journalist and student in Batumi, Georgia (Liberali.ge, 2019). The MDF found that the trolls were mobilized against critical media outlets and journalists. In particular, the report states that the Facebook "posts/comments were directed against specific media outlets as well as specific journalists. Rustavi-2 and TV Pirveli were the targets of attacks in this regard and in individual publications, also the Liberali online edition" (Liberali.ge, 2019).

Misinformation has actively targeted foreigners: on September 22, the online news outlets alia.ge, geotimes.ge and dainteresdit.ge shared a Facebook status alleging that the Vashlijvari Exaltation Church had ceased to ring its bells upon the request of Iranian residents in Tbilisi. The news articles were reportedly shared thousands of times and led to heated discussions and hateful comments in the websites' comment sections and on social

media. Dean Giorgi Sakhvadze of the Vashlijvari Exaltation Church denied the allegations and stated the church had not been involved in any confrontation on grounds of ethnicity or religion (Chimakadze, 2018). Once the netgazeti online newspaper reported the news was false, the Facebook user who had initially posted the story subsequently deleted her post but reportedly the online news outlets that shared the story did not correct the false claims in their stories (Chimakadze, 2018). Similarly, a doctored image of a street sign was shared by Facebook pages in March 2018 claiming Ivane Machabli Street in Tbilisi had been renamed 'Iran Street' (Kokoshvili, 2018). The online news sites alie.ge, infonews.ge and guriismoambe.ge shared this story and none of the stories provided links or evidence of the report (Gugulashvili, 2019).

Armenians, who are a minority group in Georgia, have also been targeted by online news outlets and Facebook pages. For instance, this story was posted by the Iberian Unity Facebook page in August 2017, which was shared over 3,900 times and reposted by Info9 and Rezonansi in March 2019 (Pertaia, 2019). Another example of xenophobic fake news was a story published by the website intermedia.ge that, according to a United Nations report, alleged that the Georgian ethnos was disappearing quicker than any other in the world; it stated that the combination of slowing birth rates and more foreigners was responsible for this development. One fake site called theguarian.com – whose design was identical to the online British newspaper theguardian.com – shared a story of a fabricated interview with a British foreign service agent who allegedly stated that, after the former president Mikheil Saakashvili was brought to power, the British foreign office launched a 3-stage plan to dissolve Russian military influence in Georgia (mediachecker.ge, 2017).

The LGBTQ community are frequent targets of online misinformation, too, which is largely peddled by online news websites and the Facebook pages of ultra-nationalist activists. For instance, on 3 September 2018, the leader of the ultra-nationalist group Georgian March, Sandro Bregadze, made a misleading statement that a gay pride event would take place at the national football stadium, Dinamo Stadium, on 9 September, which online news outlets Alia, Metronome and Kartuli Azri shared without verification. In reality, on 9 September, Georgia and Latvia were scheduled to play a football match at the stadium and – in support of the Georgian footballer, Guram Kashia, who had been a public target of ultra-nationalist and homophobic groups for wearing LGBTQ armbands during matches – LGBTQ groups announced that they would attend the match with armbands and display banners saying "#guramshentanvart", Georgian for "Guram we are with you" (Mythdetector, 2018). Only a few days later, on 7 September, the online outlets, alia.ge and resonancedaily.com, published Bregadze's Facebook status in which he falsely stated that a lawsuit by LGBT organizations and the ombudsman would legalize same sex marriages, as well as the adoption of children by same sex couples in the near future. Bregadze blamed this development on "Kashia's LGBT armband and his support!!!" while adding, "Now you can celebrate a victory over the Kazakh team manned by shepherds or the Andorran team manned by barbers and fishers!!!" (Mandaria, 2018).

GERMANY

The *Freedom on the Net* report by Freedom House places Germany above the EU average, with an Internet penetration rate of 84% and well-developed networks and ICTs (information and communication technologies) – although there is a 4% gender gap in terms of Internet access. Generally, the Internet is considered free, though a law passed by the Bundestag in June 2017, which came into full effect in January 2018, is considered by some as an infringement on freedom of speech and could lead to improper censorship of online content posted by private individuals. This law, called the Network Enforcement Act, compels social media platforms with more than 2 million registered users (excluding messenger and chat apps' users) in Germany to delete language which violates the constitution as illegal speech, such as defamation, libel, or slander, within 24 hours of being reported; and to remove content which appears to be illegal hate speech within 7 days. Companies who do not comply with the law face fines of up to €50 million.

There are two further laws in Germany which allow for content removal and targeted state surveillance. The first one is an amendment to the German Criminal Code of Procedure allowing law enforcement to install "State Trojans" to read suspects' encrypted messages; the second is the Data Retention Law which allows law enforcement and intelligence services to collect and store data of citizens' online activity on preventative grounds from 4 weeks (mainly location data of mobile phone connections) to up to 10 weeks (mainly numbers, dates and times of phone calls, messages and Internet connections). The latter has been especially criticized for not being in line with EU law based on a 2014 European Court of Justice guiding interpretation. In general, however, it seems the focus of the Cyber and Information Space structure within intelligence and the military is more on classic cybersecurity issues like network security or counterintelligence and less on information warfare and citizen surveillance. There are no reports that the country is creating significant capacity in this area. Additionally, in February 2019, the Administrative Court of Cologne ruled that the German domestic intelligence service – the Office for the Protection of the Constitution (BfV) – is not allowed to declare the entire AfD as a case for heightened scrutiny as it would "convey a negative impression to the public". The AfD had previously claimed the declaration as politically motivated, and subsequently went on to legally challenge it before the court in Cologne.

Nonetheless, state actors are generally not engaged in blocking or filtering content online. In 2015, the Federal Court of Justice ruled that blocking websites could be used as a last resort provided it was the only way to end copyright infringements. In February 2018, the first case of blocking was enforced by the Munich regional court when it compelled Vodafone to block kinox.to for uploading illegally owned content. Most other instances of removal online relate to search engine results, rather than to content as such.

In relation to misinformation campaigns and elections, an inquiry showed that no such campaign seems to have significantly influenced the election results in 2017. Nonetheless, fake news and conspiracy stories made up nearly 20% of Twitter news content during campaign time. However, microtargeting or dark advertising on social media is not a common phenomenon due to Germany's electoral and data protection laws, and because most German parties are comparatively slow in adopting a digital campaigning strategies. 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. However, the extent to which it was used was much smaller compared to, for example, the United States, as social media in general plays a subordinate role in news consumption in Germany. Nevertheless, Facebook is teaming up with the German news agency DPA to fight fake news through fact-checking. Facebook is aiming to increase the number of people fighting fake news globally from 20,000 to 30,000 by the end of 2019. Within Germany, current staff numbers are at about 2,000, and Facebook is planning to train up to 100,000 students in media literacy. Moreover, several politicians have called for a crackdown on social media bots after MPs were flooded with messages on social media and via email during the debate on the UN migration pact in late 2018. It seems as if over one quarter of the messages and tweets were from bots. Reportedly in Germany, you can buy around 1,000 bots for under €10. In addition, a 20-year-old hacker from Germany stole and leaked the personal data (identity documents, personal communication, etc.) in December 2018 as an advent calendar with new links acting as doors to more information. He cooperated with authorities and said he acted alone after being annoyed by statements from the politicians affected.

Misinformation from foreign countries, mainly Russia, is also increasingly a concern for the government. Russia has established channels such as RT and Redfish, which are especially visible on Facebook and YouTube (Figure 7). The main focus of these channels are right-wing topics supporting the AfD party or Pegida movement. However, according to reports in late 2018, Russian controlled news sites seem to be taking up left-wing and green topics too. This development has led to demands from politicians across parties to increase the government's fight against misinformation to ensure that the public decision-making process is not affected, especially in the light of the upcoming EU election. These channels present themselves as a new form of grass-roots' journalism for anybody fed up with mainstream media. Even politicians themselves are sometimes unaware of the background of such channels: Green Party MP Canan Bayram, for example, gave an interview to Redfish and was only made aware of their connection to Russia later (Figures 8 and 9, example posts). Two politicians have also given interviews to RT Germany: in 2017, Secretary of State Sigmar Gabriel and, in April 2019, Katarina Barley, Federal Minister of Justice and 'Spitzenkandidat' for the Social Democrats in the lead up to the EU election – the latter causing great irritation.

In terms of social media use by political actors, most political parties have official social media accounts. The AfD in particular seems to be quite successful online, continuously scoring high engagement numbers. However, it looks as if this success is partly due to fake accounts: in early 2019, a local AfD account was caught in what has been dubbed a 'fake-account-fail' where they commented favourably on one of their own posts (Figure 10). Other users pointed this out and the comment was quickly deleted. What probably happened is that the administrator of the AfD page forgot to change profiles when commenting on the post.

Figure 7: Number of followers of the three main channels owned by Russia

Die Kanäle von Ruptly, Redfish und "In the Now" erreichen teilweise Millionen von Menschen. (Quelle: Statista)

Source: https://www.t-online.de/nachrichten/deutschland/gesellschaft/id_84640062/russlands-medienzentrale-in-berlin-der-informationskrieg-ist-fuer-rechtsstaaten-ein-problem-.html

Figure 8: Example Twitter post on Redfish, owned by Russia (post fails to mention the intense resistance to the police the suspect exhibited before the video starts)

redfish
@redfishstream

Berlin police savagely beat an unarmed black man because of a "suspected bicycle theft". One officer takes out his gun. Passers-by get pepper-sprayed when they try to stop the violence, police call the public's bravery "violent resistance". #Kotti #DankePolizei

♡ 619 9:52 AM - Sep 28, 2018

1,030 people are talking about this

Figure 9: Example Facebook post of In the NOW, owned by Russia

Figure 10: AfD's 'fake-account-fail' (comment by Andre Wolf reads: The Dilemma of forgetting to change to your Fake-Account on Facebook before you praise yourself. A small social media lecture.)

GREECE

Over the past few years, computational propaganda and information manipulation online have swiftly moved into Greek public life. Given the politically and financially unstable situation of the country in the recent past, fake news and conspiracy online have found fertile ground in Greece. About 71% of Greek internet users are now using social media as their main source for news, followed by television (67%). Freedom House reports that Greece's parliamentary democracy is characterised by vigorous competition between political parties who are using social media for political campaigning. Recently, there have been several political controversies in Greece based on fake news. Some experts even claim that social media and fake news are so influential in Greece that they are starting to distort democracy.

The current media and disinformation environment in Greece must be viewed through the lens of the broader media ecosystem. Prime Minister Tsipras had pledged to support modern and independent news however many traditional media oligarchs are still in place. Some traditional media oligarchs have been accused by the SYRIZA party as being highly biased, and have taken strong positions against the government during the 2015 referendum in Greece. At the same time, new "Tsipras-era" media outlets are establishing themselves and continue to control the main media outlets. They have also been accused of co-opting national broadcasting to promote positions by appointing friendly journalists. In 2016 the government held auctions for television licenses, which were officially overseen by the independent National Council for Radio and Television. However, critics have accused the government of using the procedure to alter the media landscape in their favour. In 2016, the auction was declared unconstitutional, but the government has continued to pursue it.

Prior to the European elections, Greek media were worried about fake news interfering with the elections, as well as Russian trolling. While some European countries (e.g. Germany and France) passed domestic laws against fake news (which are not without controversy), Greece is keeping with EU coordinated measures such as National Election Networks to support the national election processes across the EU.

There is evidence to suggest that Greek political parties have used computational propaganda as part of their digital campaign strategy. Some reports have suggested that major parties have "dealing rooms" where people sit and coordinate the dispersal of news and texts to influence voters, order trolls and block certain news (Karaisaki, 2019). These activities mainly focus on individuals who are undecided and did not want to vote. One early example is the Truth Team (Ομαδα Αληθειας), which was set up in 2012 and was working as an informal communications machine for the right-wing government at the time and then opposition (New Democracy party).

Some evidence suggests that Greek parties have used bots to amplify the party position. Similarly, trolls and human-operated fake accounts have been active in Greece. During the Rhodes 2018 local elections, both politicians and private individuals were targeted with hate campaigns by trolls. While initially the trolls worked individually or in small units, they

quickly started coordinating and formed troll farms. Organisers created webpages where they would leave directions as to whom should be targeted and state political goals.

There is very limited work on the exact size and operation of cyber troops in Greece, or the impact computational propaganda has on public life and conversations about politics. One study by the Civic Information Office monitored conversations about Greek politics in the EU election (Papaevagelou 2019). Another study by Crisis Monitor analysed a total of 3.868 fake news mentions, which were published between 1-10 March 2019, with Twitter having the largest volume (figure 11). Thus, more people are using the language of "fake news" to discuss Greek politics. However more independent research needs to be done to understand the impact of computational propaganda in Greece.

Greek society has become increasingly polarized and these growing divisions have been key issues discussed in mis-and-disinformation, as well as conspiracy theory. Nationalist and religious groups in Greece are sometimes the source of conspiracy and disinformation. Issues pertaining to North Macedonia and territorial disputes with Turkey are also key issues. Corruption is also a major topic for mis-and-disinformation.

In addition to conversations in groups on Facebook or websites, media oligarchies are a source of misinformation. For example, To Vima – a previously well-respected media outlet – reported the US ambassador and the PM of North Macedonia were on a secret meeting in Halkidiki, a touristic northern part of Greece while in reality they were just on holidays in close by areas at the same time (ToBHMA 2019). Although the media outlet used a misleading headline, the body of the text included more factual information about the events. This story was also picked up by other media outlets including Vima.gr and in.gr.

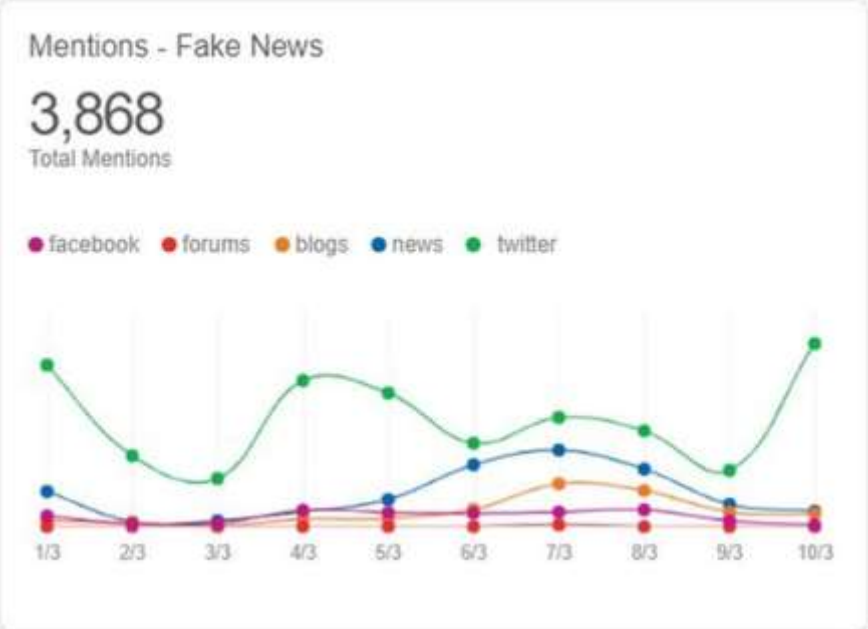
In other cases, politicians and parties are a source of mis-and-disinformation. For example, the leader of the new populist party – the Greek Solution – has claimed to have hand written letters from Jesus Christ. In 2018, a former SYRIZA communication group member was appointed as General Secretary of Information and Communication. The party's communication group was engaged in spreading false news and carrying out personal attacks on social media, targeting political opponents and producing fake news during the 2012 national election.

There have been efforts by public broadcasters to reduce the amount of hate and hate speech in Greek media. On national level, public broadcasters announced in April 2019 that they would no longer feature materials from the right wing, Nazi ideology party Golden Dawn. In an open letter the labour union of Hellenic radio and television, ΠΟΣΠΕΡΤ, stated that if the public broadcaster ERT were to include Golden Dawn content they would transform themselves "into a means of spreading an ideology that has as its core the racist and mischievous proposal of Nazism".

Finally, Greece is preparing for national elections in 2019, however, it is not clear if and when they will be held. Latest reports say the election will happen in October 2019. In the meantime, Greece is preparing for messy campaigning. National issues, political controversies and personal vendettas are rapidly taking on uncontrollable dimensions as

the elections are nearing. These issues are further amplified and complicated by the spread of computational propaganda.

Figure 11: Fake news mentions per platform



Source: Crisis Monitor, 12/03/2019, https://www.crisismonitor.gr/2019/03/12/analytics-ekloges-plisiazoyn-fake-news-bots-kai-trolls-sto-proskinio/

HUNGARY

Hungary, whose government is currently lead by Prime Minister Victor Orbán, is considered to be the only partly free democracy in the European Union. Orbán commands a majority parliament through the Fidesz–KDNP coalition, which has been criticized for dismantling Hungary's democratic institutions as well as its independent media (Freedom House – Hungary, 2019). During Orbán's nine-year rule, there has been a decline in the number of independent media outlets; for instance, 2016 saw the closure of *Népszabadság*, the largest, independent daily newspaper (Freedom House – Hungary, 2019). Many national, regional and local media have either closed or come to be controlled by oligarchs with ties to Orbán. The most striking development was the consolidation of 476 media outlets in the Central European Press and Media Foundation, whose chairman was formerly a legislator in the Fidesz party (Besser, 2019; Bognar, 2019). Critics have dubbed the Central European Press and Media Foundation a pro-government media conglomerate, controlling various newspapers, radio stations and websites. One critic, Ms Komuves, characterized it as a "media empire" and argued that "fake news and misinformation is coming from the state-sponsored media itself" (Besser, 2019).

While many still get their news via television (65%), more get their news online (85%), mostly on computers (63%) and smartphones (59%), according to the 2019 Reuters *Digital News Report*. The report details that the political climate negatively affects trust in news, which is very low (28%) and therefore online news outlets and social media networks are

widely used to access news – in fact, some 62% of survey participants stated they get their news on Facebook (Bognar, 2019). Facebook has been a popular platform for the spread of misinformation by political parties and news outlets. According to an investigation by 444.hu, an online political news site, the government employs troll networks to share and engage with pro-Fidesz content (444.hu, 2018). An investigation by the magazine *HVG* revealed that a video posted by the major news website Origo, which showed two dark-skinned men attacking a white-skinned woman in a church and was accompanied by a soundtrack of shouting in Arabic and a caption reading "Europe, 2017. Do you want this?", had actually taken place in Nebraska, USA in 2015 (Graham-Harrison & Shaun Walker, 2019). *HVG* also revealed that pro-government news outlets had posted and shared political content via their Facebook pages before the parliamentary elections in April 2018. For instance, while the page of TV2 reportedly posted no video content in the months before and after the election, in the month of the election it posted over 120 videos and its page views rocketed to over 13 million views (Graham-Harrison & Walker, 2019). According to two journalists at *Mérték Media Monitor*, online news outlets that are critical of the government receive less revenue from state advertising or firms owned by government-supporting oligarchs (Urbán and Bátorfy, 2018).

The government, too, has been criticized for posting misinformation on its official Facebook page. For instance, in 2018, the government's Facebook page posted a video that, in English, attacked Guy Verhofstadt, the Chief Brexit Negotiator for the European Union – however, the statements and images of Verhofstadt used in fact dated from 2014 (Graham-Harrison & Walker, 2019). Facebook did not remove the video, despite official complaints. As the *Guardian* suggests, the choice of English language (with an American accent) indicates that the video's target audience was probably not Hungarian. Orbán has come under steady criticism by European leaders for spreading fake news via traditional channels as well as on social media. In December 2018, as the leaders of all EU member states agreed to increase the budget of the EU's online disinformation monitoring service (EEAS) from €1.9 million in 2018 to €5 million in 2019, Juncker singled out Orbán as "the origin of fake news" in his country (Reuters, 2019). In the run-up to the European Elections in May 2019, Orbán was criticized for spreading fake news on Facebook. Specifically, the government posted campaign posters on Facebook showing Jean-Claude Junker, President of the European Commission, with George Soros, captioned "You too have a right to know what Brussels is preparing!" (BBC, 2019). This poster was particularly targeted at the European Commission's scheme to redistribute asylum seekers, of which policy the Hungarian, Czech and Polish governments have been vocal critics as this would "threaten Christian culture" (BBC, 2019). A spokesperson for Orbán, Zoltan Kovacs, defended the poster campaign against the EU's alleged plans to introduce mandatory migrant settlement quotas, stating, "Brussels continues to want to support illegal immigration, which is something the Hungarian people must know about" (BBC, 2019). The European Commission rejected this statement as "fake news" and "ludicrous conspiracy theories" (BBC, 2019).

While Facebook set up a special team to monitor the European elections, by partnering with fact-checking organizations and creating "whitelists" of reputable news outlets that would be permitted to appear in news feeds, Facebook came under fire for failing to partner with a domestic fact-checking organization in Hungary and for failing to recognize Hungarian

news outlets on its whitelists (Graham-Harrison & Shaun Walker, 2019). Facebook responded that it had not identified credible partners, which was an unsatisfactory response for many journalists concerned with the spread of government-backed misinformation on the platform (Graham, 2019). Facebook has also been criticized by the pro-government think tanks and media outlets for politicizing the platform. The newspaper *Magyar Nemzet* criticized Facebook before the European elections for becoming a political actor, and the think tank Századvég published a report in April 2019 criticizing Facebook's "principles of political correctness" which had made Hungarian politicians and public figures "victims of censorship" (Graham-Harrison & Shaun Walker, 2019). According to a poll by the think tank, 79% of Hungarians find it unacceptable that social media, such as Facebook, can "delete content based on its own political views", voicing anger at the suspension or banning of content and users for sharing anti-immigrant content.

The government and pro-government media have sustained fierce attacks on George Soros (a Hungarian–American billionaire), globalists and liberal elites, and refugees as enemies of the Hungarian people and state. For instance, the Hungarian government rejected the 2019 Freedom House report as part of the "empire" of George Soros (Simon, 2019). According to Komuves, the pro-government media peddle conspiracy theories and anti-immigration propaganda via their local, regional and national channels, which "you would otherwise get in a fringe fake news site" (Besser, 2019). The government's anti-immigrant and anti-Soros rhetoric is not just communicated via media but also codified in law. In 2018, the government introduced "Stop Soros" laws, which criminalized the provision of assistance to asylum seekers by Hungarian nationals. Furthermore, during his campaign trail for the 2019 European elections, Orbán criticized "the interference of that global, liberal mafia … players outside Hungary, manipulating huge funds, seeking to wage a campaign and interfere with the Hungarian elections" and argued that "Europe's borders must be protected against the invasion of migrants" (Besser, 2019).

Google has also come under criticism for initially granting the New Wave Media group a financial award under its Digital News Innovation Fund, which is designed to "help journalism thrive in the digital age" (Bayer, 2019). However, as the publisher of news sites such as Origo, the New Wave Media group has been criticized by journalists and researchers for publishing fake news. According to critics, Origo is a vehicle for government propaganda and a major recipient of government advertising (Bayer, 2019). Gábor Polyák, head of Hungarian watchdog Mérték Media Monitor, described Origo as "an emblematic player of the Fidesz propaganda media", which spreads "thousands of pieces of news about migrants in an extremely negative context and accompanied by false videos and photos" (Bayer, 2019). According to Politico, Origo has repeatedly been found guilty by judges of incorrectly portraying facts about government critics. In response, Google subsequently withdrew the grant given to the New Wave Media group.

Introduction

On 14 February 2019, a terrorist attack by Pakistan-based terrorist organization Jaish-e-Muhammad in Pulwama district of Kashmir, killed 40 Indian soldiers and triggered a wave of online disinformation. Within 24 hours of the attack, a doctored image of opposition Indian National Congress (INC) party leader Rahul Gandhi standing next to the suicide

bomber was debunked by the Indian fact-checking site Boom Live (Poynter, 2019). The Hindi text accompanying the photo questioned whether the INC party was involved – a deliberate attempt at using the attack to incite political tensions (AFP Fact Check, 2019a). Despite best efforts to counter the spread of false and misleading content, disinformation filtered through to credible news outlets, with mainstream channels in India and Pakistan publishing news stories that amplified rumours and misinformation about the attack (AFP Fact Check, 2019b).

The geopolitical incident prompted major concerns, as doctored, misleading and outdated images and videos circulated on social media platforms and reached millions of additional viewers through mainstream media. This could not have come at a more sensitive time: India's 900 million eligible voters – including 340 million Facebook users and 230 million WhatsApp users – will take part in India's general election, the largest exercise of democracy in history.

This short memo provides a brief background on "cyber troop" activity in India. "Cyber troops" are defined as government or political party actors tasked with manipulating public opinion online (Bradshaw & Howard, 2017, 2018). This memo provides an overview of the various tools and techniques used to amplify the spread of "junk news" on social media and suppress the voice and participation of political opponents or vulnerable populations online. Drawing on data from the Global Cyber Troops Annual Inventory,[2] we provide information about the capacity and resources that have been invested into social media manipulation in the lead-up to the 2019 general election.

We found that cyber troop capacity has grown significantly in the lead-up to the 2019 general election. While there were only a few actors involved in social media manipulation in 2017, political parties are now working with a wider-range of actors including private firms, volunteer networks, and social media influencers to shape public opinion over social media. At the same time, more sophisticated and innovative tools are being used to target, tailor, and refine messaging strategies including data analytics, targeted advertisements, and automation on platforms such as WhatsApp. In previous years, we have measured the capacity of cyber troop activity in India as being low. However, the increasing amount of money being spent on growing team sizes, advertising campaigns, and hiring private firms combined with the application of a variety of more sophisticated computational techniques underscores the growing capacity of cyber troops operating in India.

[2] The Global Cyber Troops Inventory uses a three-pronged methodological approach to identifying instances of social media manipulation by government actors. This involves (1) a content analysis of news article reporting; (2) a secondary literature review; and (3) expert consultations. More information about the approach can be found in Bradshaw and Howard (2017, 2018).

The disinformation following February's terrorist attack in Kashmir served as a warning to social media platforms, citizens, and politicians ahead of the general election. The low barriers to entry, availability of resources, and low levels of regulation on networks such as Facebook, Twitter and WhatsApp provide ample opportunities for propaganda. And an enormous population comprising a variety of castes, religions and languages, with varying levels of digital literacy, provides a fertile ground for disinformation. This, combined with the fact that political parties are deliberately exploiting these vulnerabilities as part of campaign strategies, has meant that disinformation is playing a key role in the 2019 general election.

INDIA

India has a long history of political parties using social media for political campaigning. The two main political parties, incumbent Prime Minister Narendra Modi's Bharatiya Janata Party (BJP), and opposition Indian National Congress (INC) party, both have 'IT cells' that are known to use automation, trolling and disinformation techniques. These IT cells have existed since the early days of social media, with the BJP's IT cell founded in 2007.

Political parties in India have also been known to work with private firms. Cambridge Analytica "worked extensively in India" according to whistle-blower Christopher Wylie (CNBC, 2018). The Indian IT firm Silver Touch was responsible for building Modi's NaMo app, and was linked to fake Facebook accounts (Facebook, 2019). Influencers on social media platforms are increasingly used to amplify political messages more organically to a wider audience. For example, Delhi marketing firm OMLogic Consulting has worked for both the BJP and INC to utilize the power of YouTube and Instagram influencers (*The Economic Times*, 2019a).

Although several countries have experienced external interference, cyber troop activity in India is predominantly of domestic origin. The deliberate spread of disinformation by politicians and political parties has often led to misinformation – the accidental spread of false content – as a result of hyper-connectivity and digital illiteracy. However, there have been a few cases of foreign interference by countries such as Pakistan who set up a series of fake accounts and Facebook pages about issues to do with India's general election (Reuters, 2019). Following the Kashmir attack, individuals linked to the Pakistan Army used Facebook and Instagram accounts to inflame tensions with India and push claims over Kashmir (DFRLab, 2019).

Cyber troops in India use a variety of strategies, tools and tactics to spread disinformation and manipulate public discussions about politics online. Disinformation often originates from non-credible news outlets or fake social media accounts, but disinformation is prolific in India as it also originates from mainstream media, politicians, and as part of official election strategies. *The Economic Times* and *India Today*, which has its own fact-checking project, published – both in print and in a video – a photo that allegedly showed the February terrorist attacker in a combat uniform; however, in reality it originated from an

unknown source on Twitter and was determined as fake (Poynter, 2019). Boom Live claim that political parties have "begun building teams for the specific purpose of pushing out a huge volume of propaganda and disinformation" (The Atlantic, 2018). Both the BJP and INC accuse each other of propagating "fake news" while denying they do so themselves (Reuters, 2018). And Amit Malviya, head of the BJP's IT cell, publicly acknowledged that there was "some scope for misinformation" during the election (Huffington Post, 2019).

Automation is used by political actors in India to create inorganic popularity around an individual, organization, or message. During the 2014 general election, the BJP were accused of paying to artificially boost their popularity on social media. On Twitter, Prime Minister Narendra Modi is second only to United States President Donald Trump as the most followed politician, with 45.9 million followers; however, a study by Twiplomacy (2018) claimed that as many as 60 percent come from fake accounts. There is also evidence of active networks of Twitter bots that are already being deployed ahead of the election to boost Modi's popularity. In February 2019, the hashtag #TNwelcomesModi received 777,000 mentions over two days, in reference to Modi's visit to Tamil Nadu, a southern Indian state. In response, #GoBackModi was mentioned 447,000 times by INC-supporting accounts (Quartz, 2019; DFRLab, 2019). Despite the high levels of automation on Twitter, this activity did not reach very many people, as the unsophisticated fake accounts had few followers.

Trolling tactics have also been used to suppress the political speech and participation of dissenting opinions. In the book *I am a Troll*, Indian journalist Swati Chaturvedi details the creation of the BJP's IT cell, also known as the 'BJP troll army,' which was formed in 2007 by Prodyut Bora to smear and threaten opponents online (Huffington Post, 2018). Today, around 300 workers use "strategies meant to inflame sectarian differences, malign the Muslim minority, and portray Modi as saviour of the Hindus" (Bloomberg, 2018). These attacks vary in their sophistication: from crudely automated criticism, such as #GoBackModi, to highly personalized attacks on individuals. They specifically target political opponents and journalists – especially prominent female figures – with sexual harassment and abuse. Sometimes individuals are also threatened with real-life physical attacks by online trolls (Reuters, 2018). For example, the Office of the United Nations Commissioner for Human Rights called for the government to protect journalist Rana Ayyub, after her face was superimposed on pornographic clips and she received rape and murder threats, following false quotes attributed to her on social media (Guardian, 2018). While parties explicitly deny supporting online trolls, these accounts are often aligned with party agendas and the leaders provide tacit support. For example, Prime Minister Modi follows known troll accounts on Twitter, and drew criticism for hosting 150 social media influencers at his residence in 2015, many of whom who used sexual slurs to harass women online (Guardian, 2018).

In the 2018 Global Cyber Troops Inventory, we found that chat applications were an important platform for spreading disinformation about politics. This is especially true in India: at least 50,000 election-related WhatsApp groups were created by both the BJP and INC during the May 2018 Karnataka state elections (Freedom House, 2018). The social media chief of the BJP declared 2018 the year of India's first 'WhatsApp elections', and has

reportedly “drawn up plans to have three WhatsApp groups for each of India’s 927,522 polling booths” (Time, 2019). A so-called ‘cell phone pramukh’ will operate a number of these groups and drive the party’s WhatsApp-based campaign by circulating specially designed campaign material (Hindustan Times, 2019). Parties are even using data analytics to form WhatsApp groups based on demographic and socio-economic factors, using information from the electoral roll to sort the population into groups based on factors such as caste and affluence, to achieve micro-targeted messages (Quartz, 2019).

The platform most impacted by disinformation is WhatsApp, and as a result it is increasingly scrutinized by the Indian government. Mob lynchings caused 30 deaths in India throughout 2018, which reportedly resulted directly from misinformation spread over the app – leading them to be known as ‘WhatsApp killings’ (Guardian, 2019). In one video that went viral in June 2018, footage of a child abduction was accompanied by text about ‘kidnappers’ arriving in the city to abduct children; however, it was actually a child abduction awareness video created in Pakistan. In line with their trolling tactics, political differences are exacerbated by inciting Hindu–Muslim tensions on WhatsApp. For example, right-wing Hindu groups circulated a video on WhatsApp allegedly depicting a Muslim mob attacking a Hindu woman, but in reality, it was footage of a lynching in Guatemala. Automation has also been attempted; during the state elections in 2018, the platform’s systems detected an attempt by someone in Karnataka to create dozens of WhatsApp groups in quick succession (New York Times, 2018). However, the platform has taken several steps to curb crude attempts at automating accounts or forwarding messages. These responses are detailed in the final section on government and platform responses.

In April 2019, Facebook took down 687 pages and accounts linked to the IT cell of the INC which posted about political issues, the upcoming elections, and criticism of the BJP. Facebook also suspended 15 pro-BJP pages, groups and accounts, and one pro-BJP Instagram account linked to Silver Touch. These accounts were not removed because of the content they posted, but because they engaged in “coordinated inauthentic behaviour” (Facebook, 2019).

Alongside evidence of computational propaganda on Twitter, Facebook and WhatsApp, Modi has his own app, NaMo, which launched in June 2015 and has over 10 million downloads. NaMo is a platform used by Modi to communicate with his followers. However, he has received a significant amount of criticism for bypassing traditional media channels and evading media scrutiny through its use (Bansal, 2018). And despite the Indian government putting pressure on social media platforms to control disinformation, there is a lack of content moderation on the NaMo app, making it susceptible to propaganda. One of the most prolific accounts on this app, The India Eye, was responsible for 40 percent of the 744 posts on NaMo’s default feed. Alt News – a fact-checking organization in India – uncovered extensive misinformation peddled by The India Eye on their Facebook page: at least six of the 20 most shared posts between September and November 2018 were inaccurate or misleading, exposing misinformation to its two million followers (Atlantic, 2019). Alt News discovered The India Eye had links with Silver Touch, the private firm linked to fake accounts on Facebook and Instagram. It is also claimed that Silver Touch created the NaMo app itself (The Wire, 2019). The India Eye’s Facebook page was taken

down by Facebook and is part of a wider propaganda network linked to Silver Touch (Facebook, 2019).

Networks of paid workers and volunteers disseminate sophisticated disinformation strategies across social media, responding in real time to political developments. The organization of propaganda efforts appears to be both centrally coordinated and volunteer-run. India's vast size and regional politics means that propaganda efforts are geographically coordinated. There is evidence of specific regional cells, such as the Gujarat Congress IT cell's 'Cyber Army' (DFRLab, 2019) and the BJP's 50-member team in an office in Bangalore (Boom Live, 2018). A former troll said that he was given a half-dozen Facebook accounts and eight cell phones as part of a 300-person team in a BJP IT cell (Bloomberg, 2018).

Alongside these paid workers, individuals can volunteer to assist in 'WhatsApp Group Management', being 'active on Facebook & Twitter' or 'Content Creation' among others, according to a volunteer sign-up form (Boom Live, 2018). There is a blurring of attribution between paid IT cell workers and volunteer movements. The former head of the BJP IT cell, Arvind Gupta, said in 2016 that neither the party nor IT cell had ever encouraged trolling, and that online support came from a grass-roots movement (Bloomberg, 2018). Relying on volunteers and paid workers allows the blurring of boundaries between campaigning, trolling and propaganda.

Both political parties used Facebook to target political advertisements at voters. Following the takedown of fake accounts in April, according to Facebook the INC-linked accounts spent US$39,000, and the BJP-linked accounts spent US$70,000 from 2014 to 2019 in political advertisements (Facebook, 2019). However, since February 21, when Facebook began to track political advertising, the total spending for political advertisements in India totalled 103 million rupees, approximately US$1.5 million (New York Times, 2019). There is a lack of transparency on who is amplifying political advertisements; while the top three advertisers in India were all aligned with the BJP's election agenda, none explicitly disclose their affiliation.

In response to the proliferation of disinformation, there have been a number of public and private initiatives designed to curb the spread of low-quality information online. Fact-checking has been an important response and several media organizations, such as Boom Live and Alt News, have been established to verify photos and rumours spread on social media.

Given its importance to Indian politics and everyday life, WhatsApp has received the most public criticism. Following the rumours spread on WhatsApp, the Indian IT ministry issued several warnings, stating irresponsible messages were not being "addressed adequately by WhatsApp", and that in the absence of adequate checks, WhatsApp would be considered an "abettor" of rumour propagation and subject to legal consequences (Bloomberg, 2018). In response, WhatsApp added a 'forwarded' tag to messages, limited to five the number of times a message can be forwarded, and launched an advertisement campaign giving "easy tips" to spot fake news (Guardian, 2018). Restrictions have proved ineffective, and technical tools to circumvent these restrictions are advertised to campaigners – such as one charging

a fee of 0.04 rupees ($0.0005) per message per individual, to allow a message to be forwarded thousands of times (Vice, 2019).

The day following the Kashmir terrorist attack, India's Central Reserve Police Force set up a team of 12 soldiers to fact-check social media posts (LA Times, 2019). Army Chief General Bipin Rawat said that "Our adversary will utilise social media for psychological warfare. We must also leverage social media to our advantage" (Economic Times, 2018). Given the heightened tensions with Pakistan following the Kashmir attack, the Indian army is now considering how to use social media to its strategic advantage. The defence ministry recently approved a new Information Warfare branch of the Indian army in March 2019 (Economic Times, 2019).

Battling disinformation is particularly difficult in India: dozens of languages make both automated and human moderation difficult, and the end-to-end encrypted nature of WhatsApp restricts the platform's ability to counter disinformation. Alt News even found that two of Facebook's media partners, India Today Group and Jagran Media Network, published false information about the Kashmir attack (New York Times, 2019). This demonstrates the difficulty in countering disinformation, and that they are against ingrained and institutionalized practices.

Computational propaganda efforts in India have gained increased prominence and media attention as a result of the 2019 general election. There has been a mounting body of evidence demonstrating the growing capacity of cyber troops in India to carry out social media manipulation campaigns. In the lead-up to the 2019 general election, teams have been growing, new techniques have been tried and tested, and private companies have been hired, all to give political parties a cutting edge on voting day. Given heightened geopolitical tensions, several years of social media manipulation, and institutionalized disinformation practices, these efforts are not set to disappear once the election results are announced. There is a growing recognition of the (geo)political power of social media, which could have an impact beyond the 2019 election as military investments in information warfare come to fruition over the coming years.

IRAN

There are multiple organisations reportedly tasked with social media manipulation. In September 2009, the Telecommunications Company of Iran was wholly acquired by the Islamic Revolutionary Guard Corps (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. Iran has actively engaged in cyber warfare under the IRGC, which controls the agenda of the Supreme Cyber Council established in 2012. Hacking divisions are commonly called 'kittens' and have engaged in cyber attacks and cyber espionage, by setting up personas on platforms like Facebook and LinkedIn. The true size and organisation of these groups is disputed; Ali Younesi, the former minister of intelligence and security, reported on state television in October 2004 that the Department of Disinformation had

hired thousands of agents (Library of Congress, 2012). According to Abbas Milani, an Iranian-American historian, there is an unofficial "cyber army" of 10,000 people dedicated to "cyber-fighting against enemy cultural invasions". According to the BBC, Supreme Leader Ayatollah Ali Khamenei ordered the creation of the Supreme Council of Virtual Space, and the creation of a 'cyber army' (also known as 'cyber jihadists') in 2010 (BBC, 2012). This cyber army is held to have hacked dissidents' Twitter accounts and opposition websites to redirect some users to alternative websites (Arrington, 2009). 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 aimed at discrediting Western media outlets.

Social media has long been used as a political tool for control, particularly surrounding protests. This came to public attention in the context of the 2009 Green Movement protests, when the rich and active blogosphere -- that had for a decade been indirectly political -- became explicitly used as a political tool for mobilization against the regime. Social media was dominated by pro-opposition users and reformists who shared images of the Green Movement to the outside world. Iran began systematically monitoring social media activity to demobilise protests, criminalised online activism in 2010, and passed the Computer Crime Law (Article 19, 2012). Iran has reportedly been testing a domestic 'Halal' Internet, including the national 'Mehr' (the Iranian version of YouTube) (Reporters without Borders, 2018). During the January 2018 nationwide protests, dozens of Twitter bots used tactics that ranged from calling widely shared videos of rallies fake to discouraging potential protestors from joining (BBC, 2018). Accounts were created by pro-regime users to guide protestors to the wrong locations and give the impression that the protests were on a small scale. One account posted in response to a video from a protest in Rasht, Gilan, "I just arrived here, there is nothing going on". The exact same messages by the same accounts could be seen commented on many videos between 1 and 4 January (BBC, 2018). The hashtag most associated with the events, #nationwide_protests has been used more than 470,000 times, but an analysis shows a large number of posts in favour of the demonstrations originate from Saudi Arabia (BBC, 2018).

Alongside responses to protests, computational propaganda has been used during election cycles. The presidential elections in May 2017, in which incumbent reformist President Hassan Rouhani was opposed by conservative Ebrahim Raeisi, became the object of 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. One botnet is attributed to the People's Mujahedin of Iran (Mujahedin-e Khalq, or MEK) who claim to be the Iranian government in exile, advocate the violent overthrow of the Iranian regime, engage in hijacking hashtags, flood the platform with tweets decrying the human rights record of Iran, and 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 high volume of content, for example

the pro-Rouhani account @nedayeeslahat appears to be automated, posting 52 times between 13 and 17 May, including seven posts within 40 seconds.

The MEK has an active online presence from their base in Tirana, Albania. According to The Guardian, former MEK member Hassan Heyrani said that there were several thousand accounts managed by about 1,000-1,500 MEK members in Tirana. They posted pro-Rajavi and anti-Iran propaganda in English, Farsi and Arabic on Facebook, Twitter, Telegram and newspaper comment sections, working "form morning to night with fake accounts" (Guardian, 2018). Marc Owen Jones, an academic who investigates political bots, found thousands of suspicious accounts emerged in early 2016 with 'Iran' as their location, and posted in support of Trump and the MEK. The majority of accounts tweeting the hashtags #FreeIran and #Iran_Regime_Change from December 2017 to May 2018 were created within a four-month window, suggesting automated activity (Guardian, 2018).

Telegram played a significant role in the 2017 presidential election, with both major campaigns deploying automated bot accounts to disseminate political messages (Freedom House, 2018). Conservative activists deployed a fake Rouhani bot with a very similar handle to the 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. In January 2017, it was announced that the administrators of Telegram channels with more than 5,000 members would be offered incentives to register with the Ministry of Culture and Islamic Guidance, and by April it was reported that 8,000 channels had registered (Freedom House, 2018). However, authorities moved to ban Telegram altogether in April 2018. It has an estimated 40 million users in Iran and had been subject to temporary blocks in response to protests in the past (BBC, 2018). Pavel Durov, CEO of Telegram, wrote in a blog post that the company had complied with Iranian government requests to shut down Telegram channels that called for violence during the protests (NY Times, 2018). Al Jazeera has since reported that the Iranian government has released a mobile messaging app to encourage users to stop using Telegram. The app is called Soroush, however its links to the government have raised concerns that messages could be monitored (Al Jazeera, 2018). The BBC reported that it had 5 million users by April 2018, and included a 'Death to America' emoji, alongside emojis with praise for Khamenei, and wishing death to Israel, as can be seen in Figure 1 (BBC, 2018).

The use of social media is also complemented through using traditional media outlets. The Washington Institute states that the Islamic Republic of Iran Broadcasting (IRIB) is often directly involved in the dissemination of disinformation and propaganda. It is alleged that Iran's Ministry of Intelligence and Security assists IRIB with monitoring media, and that there is an organisation within the ministry that uses psychological warfare -- known as the Department of Disinformation (Washington Institute, 2018). IRIB managed a satellite television channel called Press TV which targeted Western countries. In 2012, the channel was removed from British television for violating the Communications Act, and had its broadcasting license revoked by Ofcom, the UK regulator (The Telegraph, 2012).

Social media companies have begun to remove the accounts that have been tools in Iran's computational propaganda campaign. In August 2018, Facebook announced it had removed hundreds of Iran-based pages, groups and accounts, alleging they were part of a network linked to Iranian state media, IRIB, and taking part in "coordinated manipulation" (NBC News, 2018). Facebook announced that they had "removed 652 pages, groups and accounts for coordinated inauthentic behaviour that originated from Iran and targeted people across multiple internet services in the Middle East, Latin America, UK and US" (Facebook, 2018). They also disclosed that more than $6,000 had been spent on advertisements on Facebook and Instagram, first appearing in January 2015 and continuing as recently as August 2018. This investigation was based on a report by cybersecurity firm FireEye that identified the Iranian influence operation. They uncovered a network of inauthentic news sites and associated accounts across multiple social media platforms that promoted narratives in line with Iranian interests, such as anti-Saudi, anti-Israeli and support for US policies that were favourable to Iran. Site registration data and account-linked phone numbers from Iran, as well as the consistent pro-Iranian messaging, signals a high likelihood that this originated from Iran (FireEye, 2018). Following the FireEye investigation, Twitter and Google also suspended an expansive network of accounts and websites that had links to the IRIB (WIRED, 2018). Kent Walker, Google's senior vice president for global affairs, said they had "identified and terminated a number of accounts linked to the IRIB" which had been "sharing English-language political content in the US" (Washington Post, 2018).

Further, in August 2018 Twitter announced that it had uncovered an Iranian social media operation comprising 770 users and one million tweets. The Iranian social media accounts targeted Saudi Arabia, mentioning the phrase 'Saudi' nearly 90,000 times. The ten most frequently-used terms included geopolitical issues such as Saudi, Trump, Palestine, Israel and Syria, signalling the international rather than domestic nature of the targets (Medium, 2018). The Computational Propaganda Project analysed the tweets released by Twitter in October 2018 linked to the Iranian influence campaign. The researchers found that Arabic was the third most used language in the data set, more than 69% of the links shared were to pro-Iran Arabic-language news websites, and the most widely shared websites pushed an Iranian political narrative – including criticising Saudi Arabia and supporting Syrian President Bashar al-Assad (Elswah et al., 2019).

The Atlantic Council's Digital Forensic Lab stated that the Iranian Twitter accounts peaked earlier than the Russian troll accounts -- surging in activity in 2014 and during a smaller peak in October 2017 (Medium, 2018). Ben Nimmo, an Information Defense Fellow at the Atlantic Council, stated that the operation "was big but it was frankly clumsy," as whilst other nations' operations engage people and use sophisticated messaging, the Iranian operation was using social media to message people and amplify links to disinformation websites. Whilst there was a high number of tweets, they were "spreading fairly thinly" as the accounts simply shared links to pro-Iranian websites rather than creating human personas to engage with a community (WIRED, 2018). For example, approximately a third of the one million Iranian tweets released by Twitter contained links to AWDnews.com, part of the network of sites exposed by FireEye. This operation can therefore be seen as distinctive from other computational propaganda efforts, as it focused on web-based content, using the accounts as an amplifier for the websites rather than the source of

disinformation itself. One inauthentic news website discovered by FireEye was 'Liberty Front Press' (www.libertyfrontpress.com) which publishes political news related to the USA (the home page can be viewed in Figure 2). Liberty Front Press has also maintained social media accounts on Twitter, Facebook, Instagram, Google Plus and YouTube. Several of the Twitter accounts (see Figure 3) are linked to a phone number with a +98 (Iranian) area code and were created on the same day as each other (FireEye, 2018).

In response to the shutdown of social media accounts, Iran's foreign minister Mohammad Javad Zarif accused Twitter of double standards, by shutting down the accounts of 'real' Iranians while letting an army of fake bot accounts continue. He tweeted "How about looking at actual bots in Tirana used to prop up 'regime change' propaganda spewed out of DC? #YouAreBots". Iranian media accused the MEK, Israel and Saudi Arabia of being behind social media campaigns that have called for the overthrow of the Islamic government (Reuters, 2018). The discovery and subsequent takedown of inauthentic accounts has continued:

In October 2018, Facebook said it had identified a second influence network. It shut down 82 pages, groups and accounts, including 30 Facebook pages, 33 Instagram accounts and 3 Facebook groups, that were followed by around one million users in the US and Britain (NY Times, 2018).

In March 2019, Facebook removed 513 pages, groups and accounts for "coordinated inauthentic behaviour" with ties to Iran, operating in Egypt, India, Indonesia, Israel, Italy, Kazakhstan, and across the Middle East and North Africa. Around 1.4 million accounts followed one or more of these pages, and they spent US$15,000 on Facebook advertisements (Facebook, 2019).

In May 2019, Facebook removed 51 accounts, 36 pages, 7 groups and 3 Instagram accounts involved in "coordinated inauthentic behaviour that originated in Iran" (Facebook, 2019).
In May 2019, FireEye uncovered a network of fake American personas on accounts made between April 2018 and March 2019, which it suspects is organized in support of Iranian political interests. These accounts impersonated individuals, including Republican political candidates, to disseminate favourable messaging towards Iran (FireEye, 2019).

In May 2019, Citizen Lab discovered an Iran-aligned network of websites and online personas active since early 2016, which is used to spread false and divisive information targeting Saudi Arabia, the United States and Israel. They discovered 135 articles, 72 domains and 11 personas, that have been active since early 2016 (Citizen Lab, 2019).

In June 2019, Twitter removed nearly 4,800 accounts with ties to the Iranian government. These accounts shared global news content in line with the geostrategic views of Iran, engaged in discussions related to Israel, and targeted political and social conversations in Iran and globally (Guardian, 2019).

Figure 12: Soroush app's emojis

Source:Al Jazeera

Figure 13: Liberty Front Press news website

Source: FireEye

Figure 14: Liberty Front Press associated Twitter accounts

Source: FireEye

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 Army (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 1600 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 PM Benjamin Netanyahu. It received US$778,000 in funding (which 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 US$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 US$100m in support of "hidden propaganda" against the BDS movement and its sympathisers. The biggest expenditure (US$740,000) was reportedly budgeted to promote content on social media and search engines, including Google, Twitter, Facebook, and Instagram, while US$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 (US$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 whistleblowing letter signed by 43 serving and former 8200 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 sockpuppets 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

Generally, the Internet is freely accessible in Italy and the Italian government does not engage in any kind of censorship or blocking. The Internet penetration rate of the country is higher than the global average, but lower than the EU average. In addition, there is a north/south divide in penetration rate, with the north having a higher rate on average compared to the south. In terms of legislation there has been some controversy and criticism from international organizations and the United Nations. In November 2017 the Italian government adopted a law requiring telecommunication services to retain telephone and Internet data for up to 6 years. There was little parliamentary debate on the new legislation, the general public were unhappy about the situation and staged demonstrations. Moreover, the United Nations Human Rights Committee has raised concerns about Italian legislation relating to two issues in 2017: firstly, the fact that defamation is a criminal offence in Italy and civil libel suits against journalists and online activists continuously put great financial strain on the online media landscape of the country; and secondly, that Italian intelligence employs hacking methods and intercepts personal communications without explicit statutory authorization. However, in 2016 the Supreme Court of Italy ruled hacking by intelligence agency as constitutional. Italian politicians have subsequently tried to regulate hacking but have thus far been unsuccessful. In March 2019 the magazine *Vice* published an article about research conducted by Security Without Borders on Apps from the Google Play store which turned out to be government malware infiltrating phones and leaving them vulnerable to further hacking. According to Google, less than 1,000 people were affected, all of them Italians. The malware was likely developed by Italian company eSurv, which had been awarded over €300.000 by the Italian government to develop passive and active interception systems. Much of the incident is still unknown, however, it seems clear that the spyware is not legal.

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 relation to two crucial political events: the 2017 Constitutional Referendum and the general elections on 4 March 2018, in which M5S came first. In the hours running up to the vote, several observers registered bot activity on Twitter. This development has led to growing demand in the country that the government should find a way to effectively deal with disinformation campaigns online. However, in 2017 their attempt to pass a bill which would fine online news organizations up to €5000 for publishing false, exaggerated or biased news reports was blocked in parliament.

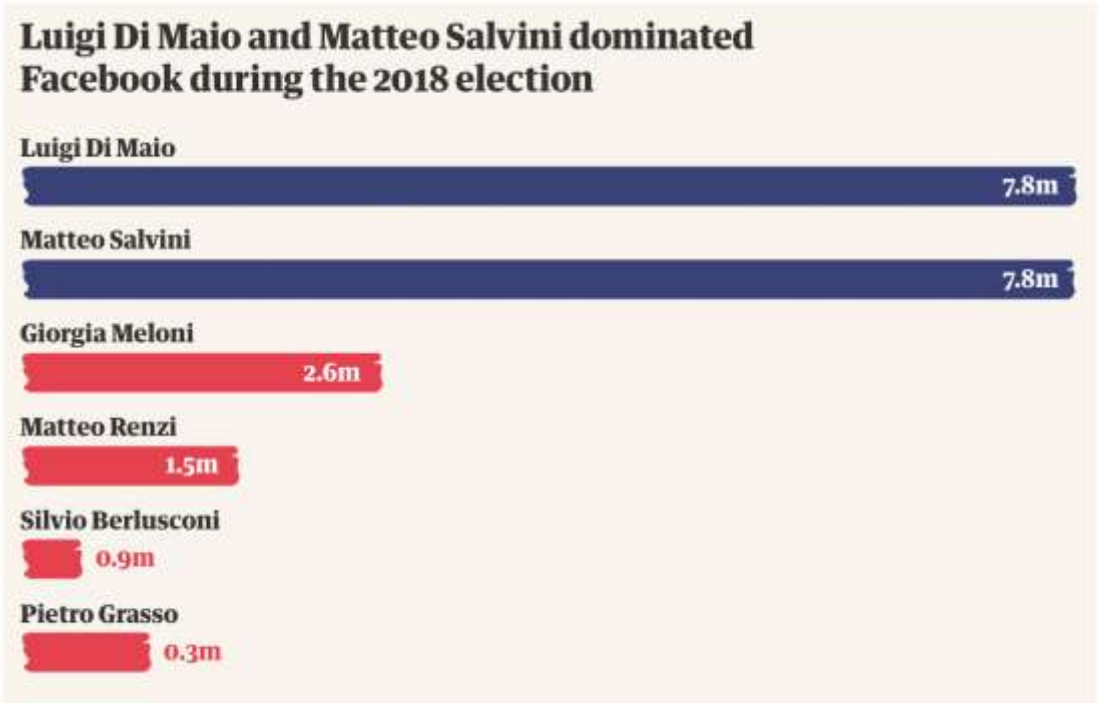
For the 2018 elections, the Northern League mobilized in particular 'voluntary' Twitter bots among its follower base, retweeting the party's official Twitter feed, @LegaSalvini, as well as employing an app that automatically embeds party postings in supporters' timelines. The M5S has for a long time 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. What is different about these voluntary Twitter bots is that these are real accounts from private individuals who turned themselves into bots, "selfbots" as the Digital Forensic Research Lab calls them, which then

all tweet the same messages (Figure 15). Normally, there is usually a “herder” or a teach which creates new accounts or repurposes hijacked accounts for their botnet. These selfbots, however, while sending out automated messages are still human as, outside these tweets, they post individual content created by the actual users.

Nearly all politicians in Italy have established a presence on major social media platforms (Facebook, Twitter, Figure 16). The populist candidates in particular were able to harness the frustration of the electorate by posting live videos on Facebook discussing issues such as migration to score high engagement numbers. Even after the election, leading politicians regularly take their debates about legislation to social media not just to comment on current issues and express their views, but also to accuse each other of political propaganda and engaging in heated debates online.

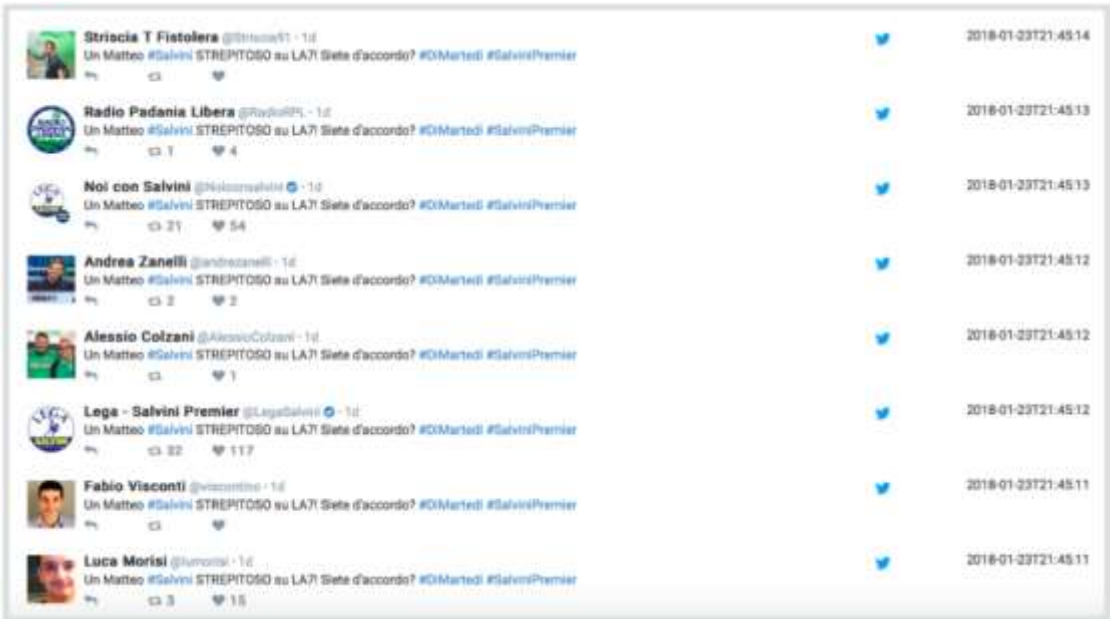
Meanwhile, the Italian government has taken official action against fake news, instituting educational initiatives in schools by adding media literacy to school curricula, and setting up a unit within the Polizia Postale (Postal and Communications Police), encouraging cooperation between ISPs (including platforms, and Facebook in particular), citizens and police to report fake news, leading to public refutations and removal requests. The project – called ‘The Red Button’ – was launched in January 2018 to allow citizens to report fake news on a portal provided by the police. The National Anti-Crime Information Center for Critical Infrastructure Protection (CNAIPIC) was tasked with analysing the reported content. There has been some criticism regarding the vague language defining fake news and the job of the CNAIPIC. Moreover, the Reuters report *Measuring the reach of “fake news” and online disinformation in Europe* relativized the impact of such sites in Italy, both in terms of average monthly reach and time spent on those websites, although the Facebook interactions of false news sites exceeded those produced by the most popular news brands. Nonetheless, misinformation and fake news remain a concern for the country: in addition to campaigns organized by domestic teams, data released in early 2018 also suggest that the same Russian company (Internet Research Agency, IRA) which was behind disinformation campaigns during the 2016 US election was also behind thousands of tweets and profiles in Italy. The US information website Fivethirtyeight.com released 9 Excel spreadsheets containing millions of tweets and profiles which US special counsel Mueller strongly suspects are from the IRA and some of the content is in Italian. So far there seems to be no official response from the Italian government. While it is unlikely that the Lega party or Five Star Movement directly supported or paid for these tweets, most of the content shared by IRA profiles was supportive of these two parties and a report from the Atlantic Council suggests close ties between both parties and several Russian individuals.

Figure 15: Facebook engagement rates by politician



Source: *Facebook data analysed at University of Pisa's MediaLab and University of Milan. Data spans 1 January to 3 March 2018 (https://www.theguardian.com/world/2018/dec/17/revealed-how-italy-populists-used-facebook-win-election-matteo-salvini-luigi-di-maio)*

Figure 16: "selfbot" network tweeting the same message at the same time



Source: *DFR Lab (https://medium.com/dfrlab/electionwatch-italys-self-made-bots-200e2e268d0e) January 25, 2018*

KENYA

Kenya has seen its fair share of fake news and misinformation campaigns during the 2017 presidential election when parties hired bloggers and paid as much as US$6 million to Cambridge Analytica to support online campaign efforts with their social media insights. Moreover, the *Freedom on the Net* report on Kenya attests that social media bots had quite

a significant influence on online discourse with 25% influence during the August 2017 election and 28% influence during the re-election in October 2017. The 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 concerning the election was spread using 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. It seems most of this activity stopped after the election, but effects continue to be felt across Africa: elections are becoming more and more expensive as candidates are no longer just politicians but represented as brands that are carefully managed by growing online campaign teams.

In terms of access, while a majority of Kenyans have phone subscriptions (91% in 2018) and access to the Internet, there is still a gender and urban–rural divide. Facebook continues to be the most popular platform with roughly 5 million daily users followed by WhatsApp. In general, Kenya does not filter or block Internet access, however, the government does regularly remove content – or requests content be removed – from platforms such as Facebook. Content to be removed is usually that which is deemed illegal in Kenya, such as projects advocating LGBTQ+ rights.

In recent years, Kenya's government has become an avid surveillant of its citizens' communications. Several state actors carry out surveillance essentially without any judicial oversight. The main intelligence agency is the National Intelligence Service (NIS), which is responsible for both national security and foreign intelligence. The NIS has direct access (meaning they can access data without prior notice or judicial authorization) to Kenya's telecommunication network and Internet providers. There remains some hope for privacy, as Freedom House reports that in April 2018 the High Court of Kenya ruled as unconstitutional the Device Management System which directly accessed mobile subscriber data. Nonetheless, spending for defence and intelligence forces continue to rise; from 2016 to 2017 the budget rose from 98 billion KSh (£738 million) to 124 billion KHs (£959 million). Additionally, the country further receives significant sums of money to support their counterterrorism efforts from several Western countries, including the US and the UK.

In addition to the NIS, the Kenyan police services also have a surveillance mandate, allowing them to collect information about serious crimes, including cybercrime. In 2012 the Kenyan Communication Commission (state-owned corporation) announced the establishment of a system allowing authorities to monitor incoming and outgoing digital communications. All Internet Service Providers were requested to cooperate as the Commission deemed this step necessary due to a continued rise in cybercrime. In late 2016 the Communication Authority (CA) (governmental regulatory body of Kenya's communication sector) finalized a contract

with the private Israeli web intelligence company webintPro to use their software in future projects. An article published by MintPress News (an independent journalism watchdog) on private Israeli spy companies meddling with elections around the globe mentioned that such companies and their respective spyware are suspected to have been used in Kenya, although there are no official statements. In January 2017 the CA announced further measures costing a total of 2 billion KHs (£15.2 million) to monitor communication and communication devices in Kenya. One measure was the Device Management System, which the High Court has deemed unconstitutional; a second is a social media monitoring project for which spyware provided by webintPro is used to capture and analyse open-source data, particularly that from social media. This enables real-time surveillance of individual users.

In February 2019 it was announced that Kenya would officially join an international coalition against Islamic State (IS), hoping to benefit from shared intelligence. Their efforts in collaboration with the coalition will focus on preventing the flow of foreign terrorist fighters across borders, supporting the restoration and stabilization of essential public services in areas liberated from IS, and fighting against IS's ongoing propaganda. This latter pledge is one of the few mentions of counter-propaganda measures in which Kenya is engaging. In general, it seems the country is not working on psyops operations fighting propaganda as other African countries (e.g. Nigeria) allegedly are. The only information-controlling measures the Kenyan military and intelligence have been accused of are spying on journalists, political actors and activists as well as occasionally pressuring them in an effort to control public news narratives; however, this is hardly done openly or on a large scale. It appears that the authority's standard action for dealing with news agencies or activists posting information or organizing events which do not sit well with the government is to accuse them of spreading hate speech, rumours and propaganda. For example, on 14 February 2019 local news reported on a shootout between two gangs near the border of the Turkana County quoting a member of parliament: "We need security beefed up. Turkana County News social site has manufactured rumors and propaganda…". What exactly the MP was referring to is unclear: The Twitter page of Turkana County News has been inactive since mid 2018, and their website does not seem to mention the events of the shootout.

In light of the government's habit of accusing activists or news organizations of spreading rumours and propaganda, several groups – including Freedom House and the New York-based Committee to Protect Journalists – have criticized the Computer Misuse and Cybercrime Act which passed into legislation in May 2018 for repressing online liberties even further. The act imposes up to 10 years imprisonment and hefty fines for the publication of "false" or "fictitious" information that results in "panic" or is "likely to discredit the reputation of a person". In June 2018 the Bloggers Association of Kenya successfully petitioned some of the provisions of the law, which is now suspended awaiting further decisions of the High Court.

Meanwhile, social media has proved how it can be a powerful political tool, as demonstrated by one particular headline in early 2019: in February, Twitter users started trolling leaders of the Wiper Party after they commented on a protest against the governor and chairman of the Wiper Party, Kivutha Kibwana (Figure 17). His ideology and personal beliefs allegedly

clash with party positions, which is why protesters wanted him to step down. This was not the first time Kibwana had been asked to step down, but it seems this attempt finally proved successful as the politician announced his resignation as chairman of the party, revoking all ties with them. However, Wiper Party officials now say that since his announcement there has been no official communication between the party and Kibwana concerning him stepping down, so it remains to be seen how this story plays out.

Figure 17: Twitter users trolling Ndolo for speaking out against protests asking Kibwana to step down

Source: *https://www.pulselive.co.ke/news/kenyans-troll-kalonzo-musyoka-after-followers-attack-prof-kivutha-kibwana/mq80m1v (08/02/2019)*

MACEDONIA

The Republic of Northern Macedonia, formerly FYROM, has been described as a "partly free" parliamentary republic, which has a "highly polarised" media landscape and struggles with frequent intimidation and attacks against journalists (Freedom House – Macedonia, 2019). Macedonia was brought to international attention when it was revealed that teenagers in Veles, Macedonia had run a large-scale fake news campaign during the US

Presidential Election, involving over 100 pro-Trump websites and Facebook pages (Subramanian, 2017). Ahead of the 2019 European Elections, Macedonia was one of the key countries on Facebook's watchlist of countries that had conduct "coordinated inauthentic behaviour" in foreign elections (iROZHLAS, 2019). According to a Code of Conduct agreed with the European Parliament to publish monthly reports of acts taken to fight misinformation, Facebook revealed in March 2019 it has removed 212 profiles in Macedonia and Kosovo (Rdanske 2, 2019). By early may, some 2,632 pages, groups and user accounts had been removed in total, including many in Macedonia (iROZHLAS, 2019).

Domestically, while television remains the primary source of news for the bulk of the adult population, studies show that the youth stay informed online via social media networks (Kalinski, 2018). According to a survey of 1,015 respondents by the Institute for Communication Studies, more than one third of respondents spend 1-3 hours per day on Facebook and one quarter said they spend more than 3 hours per day on Facebook (Kalinski, 2018). Social media networks like Facebook are becoming increasingly popular as sources of information about news and politics. Some 71% of youth between 15 and 24 agree that "It is important for me to share the opinion of public figures that I follow on social networks". Almost half of the respondents reported they follow political news and parties (49%), entertainment news (39%) and economic news (33%) (Kalinski, 2018). The increasing reliance on news on social media has alarmed commentators who are concerned that online news as well as mis- and disinformation, which are prevalent in Macedonia's media ecosystem, are becoming increasingly popular with the citizenry, who according to a study by the Open Society Institute, Macedonia ranks 35th of 35 European countries measured by resilience to fake news, based on limited media freedom and poor literacy (Veselinovic, 2018). An investigative journalist, Sashka Cvetkovska, commented "every day we face the overproduction of fake news" and citizens have limited confidence and trust in media (Civil Media, 2018). For instance, in April 2019, the presidential elections were marked by disinformation on social media, spread by presidential candidates, political parties and online news sites. According to crithink.mk, a media literacy campaign, there were five main disinformation subjects. Firstly, if no candidate achieves 40% of the vote, the current President Gjorge Ivanov would remain president until the next election (announced on Twitter). Secondly, the presidential candidate Stevo Pendarovski hired his brother to the Intelligence Agency causing uproar about nepotism (he has worked at the IA since 1999). Thirdly, on Election Day, a sensationalist story stated a Pakistani hacker sold data from the Voters List of Northern Macedonia on a Russian hacker forum. Fourthly, on Facebook and Twitter, fake stories were shared that after the results of the first round of elections, the candidate Zaev asked the DUI to boycott the elections to win the vote in the second round. Finally, hours before the second round of voting, stories were shared falsely claiming that the candidate Pendarovski had helped fake elections in 2004 (Crithink.mk, 2019).

In September 2018, the country held a referendum to rename itself from the Former Yugoslav Republic of Macedonia (FYROM) to the Republic of Northern Macedonia in accordance with the Prespa Agreement with Greece, which if passed would start preparations for Macedonia's accession to NATO and the EU. Prime Minister Zoran Zaev was a proponent of the name change and received backing from Western nations, including

Angela Merkel of Germany, while nationalists and President Gjorge Ivanov opposed it (Metodieva, 2019). Whilst 92% of voters voted in favour of the name change, only one third of eligible voters turned out to vote, which is largely attributed to an effective boycott campaign conducted both online and offline which was supported by political groups. according to Goran Nikolovski, director of Macedonia's Security and Intelligence Service, Macedonia was subject to significant disinformation campaign on social media that sought to boycott the referendum, spreading anti-EU and anti-NATO messages supported by bots and trolls on Facebook and Twitter as well as tens of anonymous news sites (Metodieva, 2019; Veselinovic, 2018). Intelligence officials were confident the disinformation campaigns originated from Russia, which opposed Macedonia's NATO and EU membership ambitions, largely because of the campaign's strong pro-Russia rhetoric on social media. The hashtag #boycott was used on Facebook and Twitter and according to analysis by the German Marshall Fund, about 40 Facebook accounts were created every day in the weeks before the referendum with the sole purpose to amplify boycott content (Metodieva, 2019). On Twitter, #boycott was mentioned more than 24,000 times and was tweeted more than 20,000 times; a significant portion of the accounts that shared this hashtag, analysis found, had been created in August 2018 and shared the same features having a Macedonian name and a random string of numbers as usernames (Metodieva, 2019). According to the think tank Transatlantic Commission on Election Integrity, Twitter accounts less than 60 days old had made up 12.5% of all Twitter accounts in Macedonia and their overall activity by tweets and retweets had made up 10% of all Twitter activity in Macedonia (Veselinovic, 2018; Радио Слободна Европа, 2019).

MALAYSIA

Computational propaganda efforts in Malaysia are not a recent phenomenon; since the general election in 2013, political parties have paid online commentators (known in Malaysia as 'cyber troopers') to defend government policies and attack the opposition (Freedom House, 2018). Last year witnessed the fourteenth general election, which saw the then Prime Minister Najib Razak's Barisan Nasional (BN) coalition (led by the United Malays National Organisation (UMNO)), which had ruled for 61 years, defeated by the Pakatan Harapan (PH) pact. The new PH government has indicated it will review the controversial Malaysian Communications and Multimedia Commission (MCMC), which had routinely limited content on blogs and social media (Freedom House, 2018).

As early as 2008, political support by bloggers is alleged to have been a factor in the election result (Johns and Cheong, 2019). In a statement in 2011, Lim Guang Eng, current Finance Minister and the then Chief Minister of Penang, stated that a "new army of cyber troopers… is proof that the 13th general election [in 2013] will be the dirtiest election yet" (Guan Eng, 2011). The 2013 election saw a boom in social media usage, and BN advertising expenditure increased dramatically for online advertising on Facebook and Google (Leong, 2015). The prevalence and importance of computational propaganda has increased with each election and, in January 2017, UMNO "urged all its members to master the use of the social media" ahead of the 2018 elections. In March, the party called on local divisions to form IT bureaus to "counter the slander" on social media (Freedom House, 2018). The prevalence of

disinformation and political attacks online has led to Malaysians describing the online battle between BN and PH supporters as “cyber-war” (Leong, 2015; Hopkins, 2014).

The BN cyber troopers are coordinated by the New Media Unit (NMU), which is partly led by the UMNO Youth Executive Committee – the NMU is a formal unit of the UMNO Youth Wing, but also a loosely associated collection of bloggers, thereby allowing a degree of plausible deniability for unofficial activities (Hopkins, 2014). By 2013, BN reportedly had a network of 80 cyber troopers who ran thousands of fake social media accounts (WIRED, 2018). These groups have tacit government backing, and the government has made no secret of the fact that it has support from a huge network of cyber troopers, some of whom are directly paid by them (WIRED, 2018). A media consultant said the BN NMU consisted of Facebook and Twitter users alongside bloggers, and that the unit had trained more than 2,000 people in social media usage (Leong, 2015).

According to local news sources, in the run-up to the 2013 election, the Parti Rakyat Sarawak (PRS) used cyber troops to gain political support. Cyber troops are often based in New Media Units (‘Unit Media Baru’ (UMB) in Malay). The party relied on a five-member team comprising party members, selected based on their interests and wide-ranging knowledge about political issues and activity in the state. These members allegedly underwent training in Kuala Lumpur together with UMBs from other parties prior to the 2011 state election (Yap, 2012). The 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).

Alongside elections, social media has also been used during social movements such as Bersih (a movement for clean and fair elections) which took place from 2007 to 2016. In an analysis of tweets collected on Malaysian Twitter during the Bersih 3 rally, 36 users were responsible for sending 1,117 messages, with many being duplicate messages sent from different accounts, signalling a level of coordination and possibly automation. These messages used the well-documented tactics of using sock-puppet accounts for astroturfing (faking opposition to the protests) and intimidation (urging people not to take part) (Johns and Cheong, 2019).

Just weeks before the May 2018 elections, the Atlantic Council’s Digital Forensic Research Lab (DFRLab) reported that bot accounts were flooding Twitter with tens of thousands of pro-government and anti-opposition messages. The tweets included visuals illustrating Malaysian government policies and questioning opposition policies. Hashtags expressed disapproval of the PH opposition coalition including #SayNoToPH and #KalahkanPakatan (Malay for ‘Defeat Pakatan’). These were used 44,100 times by 17,600 users from April 12 to April 20, with 98% of the users appearing to be bot accounts (Ananthalakshmi, 2018). A source said Twitter had suspended 500 accounts involved in the messages on the Malaysian election, as they involved spam or malicious automation (Ananthalakshmi, 2018). The information technology bureau of UMNO, the ruling party until May 2018, said it was not behind the bots and it did not know who was (Reuters, 2018). However, 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. Further, in a social

media campaign ahead of the 2018 elections, Joe Lee, a social media consultant, launched #pualangmengundi, or "go home to vote", 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 of pro-government messages (Seiff, 2018).

There is evidence of private companies being contracted to assist in these efforts. According to the DFR investigation, nine of the top 10 most active bot accounts containing anti-opposition hashtags and pro-government messages had Russian-sounding names and used Cyrillic script. Donara Barojan, a research associate of the DFR Lab said: "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). Further, there is evidence that CA Political, an offshoot of Cambridge Analytica, supported Malaysia's BN 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).

Following the 2018 election, computational propaganda efforts have not decreased. In September 2018, the Crown Prince of Johor alleged that both he and his father, the Sultan of Johor, were being monitored, as well as having cyber troopers planted on their personal and 'Johor Southern Tigers' Facebook pages (Tan, 2018). The Crown Prince said that "there are cybertroopers planted, waiting in case there is something that does not go down well with certain higher ups" (Tan, 2018). In response, the Inspector-General of Police assured that the social media pages were not being monitored (Pei Ying, 2018). In April 2019, an assemblyman publicly claimed that he was being attacked by pro-Penang Transport Master Plan cyber troopers for his outspoken views on transport projects (MSN, 2019). While claims are often made, these can be politically motivated and unsubstantiated.

Fake news has been around in Malaysia for many years. Opposition groups use the term 'fake news' to describe regime propaganda, and the regime has used it to counter questions and critiques posed by local and international news outlets. The MCMC found in 2017 that 89% of Malaysians obtain news online, and that the top social media platforms were Facebook, WhatsApp and YouTube – which facilitated dis- and misinformation related to politics, religion, health and crime (Yatid, 2019). (However, it is worth noting that the MCMC is a government unit known for media control and suppressing dissent.) During the 2013 general election, false information spread that 40,000 Bangladeshi nationals were being brought to Malaysia to swing the votes to benefit the ruling coalition (Yatid, 2019).
In response, a 'fake news' portal called *Sebenarnya* ('the truth' in Malay) was set up in March 2017 by the MCMC to enable Malaysians to check the validity of news (Nain, 2017). Further, in April 2018 the BN government enacted the Fake News Act, claiming that the majority of Malaysians encounter fake or unverified news on WhatsApp, Facebook and blogs. The law covers news, information, data, reports, images or recordings that are wholly or partly false, and states it is an offence to possess, produce, offer or share this fake news content (Freedom House, 2018). The law was rushed through in just one week, in the weeks

preceding the general election; leading the opposition lawmaker Ong Kian Ming to tweet that it was an "attack on the press and an attempt to instil fear" (Guardian, 2018). Prior to winning the election, the now-ruling PH pledged to abolish oppressive provisions in the law (Yatid, 2019). In September 2018, the opposition-led Senate blocked an attempt to repeal the 'fake news' law by the PH (Reuters, 2018). However, in April 2019, Prime Minister Tun Dr Mahathir Mohamad said the government still intends to repeal the law (Kannan, 2019).

Blocking has also been used as a tool to curb online blogs exposing political scandals. 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 1MDB scandal (BBC News, 2016). Additionally, the MCMC periodically instructs websites to remove content, including some perceived as critical of the government (Freedom House, 2017). The MCMC claims to have taken action against 4,358 fake accounts between 2017 and 2018, including blocking 40 websites, portals and blogs that disseminated so-called fake news (New Straits Times, 2018).

MEXICO

According to Freedom House's *Freedom on the Net 2018* report, Mexico's democracy is considered relatively free with regards to Internet freedom, ranking it 40th with a score of 13. The country has suffered from an increase in physical violence, especially related to drug trafficking, and this has led to over 500 attacks against journalists, at least 4 reporters were killed in 2018 according to the *Freedom on the Net* report. It is second only to Syria in 2017 in the number of attacks against the media (Andalusia Knoll Soloff, 2018).

One particular aspect of Mexican social media usage culture stands out from the rest of Latin America: the country uses Twitter a lot, with a penetration rate of 49% of the population according to the Reuters Institute (Statista, n.d.). Many social media platforms are used during elections, including Twitter, Facebook and WhatsApp. As in many Latin American countries, Mexico also has a high penetration of WhatsApp users, and 35 million out of the 60 million Mexicans are Internet users (Argüello, n.d.).

The first incident of online disinformation was recorded in Mexico in 2012, and the use of this strategy has increased since (Freedom House, 2018). In 2018, the use of automation skyrocketed as the Mexican presidential elections approached. The country held the biggest elections of its history in 2018, with over 18,000 federal seats and 17,500 seats at regional level being disputed (Instituto Nacional Electoral, n.d.)., These were also the country's most violent elections, with a murder count of 152 politicians, 48 of which were pre-candidates and candidates running for office (Etelekt, 2018).

These Mexican elections were marked by an increase in digital campaigning expenditure, averaging 25% of campaign budgets, up from 5% six years previously. The estimates are that Andrés Manuel Lópes Obrador (commonly referred to as AMLO) spent approximately

88 million pesos, while candidates Ricardo Anaya and José Antonio Meade spent 338 million and 302 million pesos respectively (Forbes Staff, 2018).

Unlike many other cases, concerns around disinformation in the Mexican elections mostly related to the intense use of automation and fake accounts. Artificial amplification of messages was used widely, especially to target the leading (and later elected) candidate, AMLO.

Mexico has seen the rise of an interesting power struggle around disinformation. Firstly, there is the rampant increase of the "diseconomy", an expression used to describe the burgeoning disinformation industry. One famous example is the case of Carlos Merlo of Victory Lab, known as the "fake news millionaire" in Mexico (@DFRLab, 2018). He claims to have command over millions of fake accounts and hundreds of bogus websites old enough to look authentic. Though the numbers might be an overstatement, Merlo claims that he can boost online content into mainstream media, operating for fees ranging from 49,000 pesos (USD2,444) to one million pesos (USD50,000) a month.

Secondly, a crowd-sourced fact-checking initiative has been set up by a coalition of interested parties (Argüello, n.d.). Agencies and media outlets, with the financial support of tech companies such as Google and Facebook, and organizations such as the Open Society Foundation, teamed up to fight the country's surge in disinformation. The initiative began in September 2017, when Mexico suffered a terrible earthquake. Named Verificado, checkers received flagged dubious news circulating online and used their network of users to fact-check information. This initiative was replicated during the Mexican elections as Verificado 2018.

Research shows that Victory Lab is likely to have orchestrated a whole network of bots to amplify content online. These botnets were hired from countries around the world, including India, Brazil, and Russia. Even though thousands of accounts were in Russian, research has not identified any tangible link between those accounts and the Kremlin (Christopher Woody, 2018).

As mentioned before, these elections were particularly violent, and much of the aggression was fuelled by disinformation in both local and regional politics. The most emblematic case is that of the gubernatorial elections of the state of Puebla, where people were attacked – and some even murdered – and ballots stolen or burned (Alberto Melchor, 2018). The elections in Puebla registered hundreds of electoral irregularities in the dispute between candidates Miguel Barbosa and Martha Erika Alonso. Competing hashtags which claimed victory before results were announced were amplified by automated accounts. Distrust continued to worsen as the electoral court demanded a recount of the votes two and half months after election day. The battle raged on after the announcement of the official results and spiked after candidate Martha Erika Alonso and former governor Rafael Moreno Valle died in a helicopter crash.

What became clear from these elections is that disinformation had a very wide reach in Mexico, but there it is hard to tell what the impact was at the federal level. Nonetheless, in

the local and gubernatorial levels, it is clear that disinformation catalyzed violence towards citizens and politicians.

MOLDOVA

This is the first year that Moldova is featured in the cyber troops inventory, and its computational propaganda efforts appear limited and under-reported. The presence of computational propaganda in Moldova has come to light this year as a result of parliamentary elections in February 2019, and a takedown of coordinated inauthentic accounts by Facebook. This is set in the context of limited independence of traditional media outlets, with Freedom House noting the presence of disinformation and manipulation in the media (2018).

Ahead of the parliamentary elections on 24 February 2019, a network of Moldovan troll accounts set up fake Facebook accounts to pose as legitimate voters and civic groups. These accounts disseminated fake news, disinformation and memes ahead of the elections (Synovtiz and Raileanu, 2019). The fake accounts also coordinated with legitimate actors to overwhelm web forums and manipulate the online debate. The accounts typically posted about local news and political issues, but also shared manipulated photos, divisive narratives and satire. In the takedown of activity on 13 February, Facebook removed 168 accounts, 28 pages and 8 Instagram accounts for "engaging in coordinated inauthentic behaviour targeting people in Moldova" (Facebook, 2019). Facebook concluded that "some of this activity was linked to employees of the Moldovan government" (Facebook, 2019). The government responded by suggesting that any fraudulent activity by government workers on behalf of the Democratic Party was carried out by rogue workers (Synovitz and Raileanu, 2019). The accounts spent less than $20,000 on ads on Facebook and Instagram, which were paid for in US dollars, euros and Romanian leu (Facebook, 2019).

Fake accounts impersonating of both organizations and individuals are a common tactic in Moldova. For example, an organization targeted by a fake account was the fact-checking group StopFals – a misinformation watchdog based in the capital, Chisinau, and created to call out misinformation on social media (Synovitz and Raileanu, 2019). Cornelia Cozonac, the director for the Centre of Investigative Journalism in Moldova, was impersonated when trolls made a clone of her account and posted messages in her name to attempt to discredit her. These messages were republished by obscure news websites and then amplified by other trolls on social media (Necsutu, 2019).

Civil society efforts have attempted to fight back against computational propaganda in Moldova. Civic activists Vlada Ciobanu and Dumitru Alaiba launched the crowdfunding campaign Adopta un trol ('Adopt a troll') with the aim of paying individuals to expose the troll factories that they work for (CIMUSEE, 2018). Further, the application Trolless – an online tool developed to spot inauthentic accounts – compiled a database of accounts which contributed to Facebook's takedown in February 2019 (Broderick, 2019).

So-called fake news remains a widespread problem in Moldova. Anti-European Union messages often proliferate, such as the rumour that the EU was sending thousands of Syrian

refugees into Moldova (Broderick, 2019). Reports allegedly coming from Russian TV and retranslated by the Moldovan TV channels claimed that there was a massive wave of Syrian refugees on the way to Moldova, with one report claiming that an "invasion of 30,000 Syrians in Moldova" would occur if the pro-European candidate were to win the election (Polygraph, 2018). A further widespread fake news scandal was the report in May 2018, during the elections for Chisinau's mayor, that 'Chisinau will be rented to the Arab Emirates for 50 years' if the pro-European candidate Andrei Nastase was elected. StopFals debunked the TV report, as the original and untranslated Al Jazeera report never mentioned Moldova but was in fact about relations between the UAE and Yemen (Polygraph, 2018). The Digital Forensic Research Lab also found that the Romanian translation of the report, with the subtitles alleging the rental of Chisinau to the UAE, was completely inaccurate when compared to the original Al Jazeera report in Arabic. Despite this, the doctored video was viewed more than 303,000 times on Facebook alone (DFRLab, 2018).

Efforts at disinformation in Moldova are interconnected with Russia, as independent Moldovan media have been engaged in an information war with the Kremlin for years (Broderick, 2019). The Ukrainian NGO Prism, in a 2018 study on disinformation resilience in central and eastern Europe, found that Moldova was the "most exposed country in Eastern Europe to Russian propaganda" as a result of the dominance of Russian-language media, the Russian orientation of the church, and the mistrust in the political class (Necsutu, 2018).

THE NETHERLANDS

As with many established democracies, there are few reports on government-organized misinformation or online propaganda campaigns in the Netherlands. As in many other countries, Dutch political parties (especially the nationalist PVV) are active users of social media to communicate with their voters. Recent reports have focused on Geert Wilders, chairman of the PVV, as most disinformation – mainly anti-Islam content – originates from him. However, in light of the published list of candidates running in upcoming state elections, the newspaper *Dagblad van het Noord* found right-wing, anti-Islam tweets from several candidates campaigning for border control (Figure 18). For example, in September 2017, Wilders called prophet Mohammed a "paedophile, mass murderer, terrorist and madman" (Figure 19). In November 2018, the Turkish Islamic Cultural Federation (TICF) which represents 144 Dutch mosques, sent a request to Twitter asking them to shut down Wilders' Twitter account as he was violating Twitter's terms of use. The TICF said it would consider legal action should Twitter not take any action, however, since the early November 2018 request there have been no reports of them taking any further steps. The TICF's request has led right-wing press to challenge the Federation for disrespecting –amongst other things – Wilders' right to free speech even outside of the Netherlands.[3] Meanwhile, Wilders has seen his Twitter following shrink throughout 2018 as Twitter culled fake accounts. Wilders lost about 15% in summer 2018 (150.000 followers), while Prime Minister Rutte lost only a handful of followers.

[3] For example, see the report from the Austrian newspaper "Unzensuiert": https://www.unzensuriert.at/content/0028166-Niederlaendische-Moslems-wollen-Twitter-Verbot-fuer-Geert-Wilders

The Netherlands is also facing trolling and misinformation attacks which are likely to have originated from the Russian *Internet Research Agency* (IRA). The first incident, traced back to the IRA in 2016, has been dubbed the 'Dutch Terror Threat'. At that time, the EU was in the process of passing a deal with the Ukraine to strengthen economic relations for which the Netherlands held a referendum to decide their position within the EU. On 18 January 2016, a YouTube channel operated by the IRA released a video (Figure 20) showing six soldiers speaking Ukrainian, burning a Dutch flag, and threatening to conduct terrorist attacks in the Netherlands should the referendum result in the Netherlands not support the Ukraine deal. The video has since been declared fake, but the fact is that the Netherlands did reject the deal between the EU and the Ukraine, although it remains unclear how influential the video was. In general, investigations have found that Russian trolling activity in the Netherlands has not led to a 'Trump effect' and remains relatively uninfluential. Nonetheless, the numbers are impressive: Russia-based professional trolls have sent over 900 tweets in Dutch in the past 2 years, and set up more than 6,000 troll Twitter accounts posing as genuine Dutch citizens. These accounts have generated more than 30,000 messages in English and have a combined following of more than 9.5 million. Most tweets and retweets focus on conspiracy theories surrounding the MH17 disaster and Wilders. Interestingly, messages in English seem to circulate better than those in Dutch.

As a reaction to these trolling activities from Russia and misinformation campaigns from the PVV, the Dutch government decided to launch an Anti-Fake-News campaign on social media lasting 4 months during both the state elections in March and the EU parliamentary elections in May 2018. Interestingly, the lower house of the Dutch parliament had previously (March 2018) asked for the EU Commission's Anti-Fake-News Watchdog task force to be scrapped after several Dutch newspapers had been reported to it for allegedly spreading pro-Russian fake news. Multiple parties (PVV, Groen Links, DGP) have argued that the task force is meddling with the Dutch free press. It seems, though, that following concerns that the EU election may be the most hackable election across Europe, the Dutch government is rethinking their approach.

Meanwhile, the Dutch intelligence services – mainly the AIVD – are also reacting to the growing threat of digital espionage and foreign trolling. Thus far, their focus has been on Russian spies and hackers. In October 2018, they accused Russia of plotting to hack the Organisation for the Prohibition of Chemical Weapons as well as trying to interfere with the investigation of the MH17 flight from Malaysia Airlines which was shot down in the Ukraine on its way to Kuala Lumpur from Amsterdam in 2014. These accusations were made after the AIVD caught four agents from the GRU unit 26165 (Main Directorate of the General Staff of the Armed Forces of the Russian Federation, officially now G.U. but still commonly known as GRU) earlier in April 2018 and analyzed their computers. In the aftermath of this arrest, Dutch intelligence was apparently also able to reveal the identify of more than 300 Russian operatives working across Europe. Lastly, the budget received by Dutch intelligence increases year-on-year; for example, the government has upped their payment for untargeted cable tapping from €15 million to €35 million in 2019. The National Cyber Security Centrum once again highlighted that national security continues to be under threat from digital attacks in their 2018 report.

Figure 18: Bart Onnes is on the list of the PVV to run in the March State elections (English: What do you mean refugees? They're rapists and people that spread hate. Get them out.)

Figure 19: Geert Wilders calling the Prophet Mohammed a "paedophile, mass murderer, terrorist and madman"

Figure 20: Screenshot of the Dutch terror threat video

Source: https://vk.com/video-78437341_171493431?list=889d215a3d31f63806

NIGERIA

At the time of writing, Nigeria is only a few days away from presidential elections, set to take place on 23 February 2019. The two main candidates are the current president Muhammadu Buhari who has been re-nominated by the All Progressives Congress (APC). Former Vice President Atiku Abubakar, People's Democratic Party (PDP), is the main opposition candidate. The main issues on which campaigns are focusing are tackling economic growth, corruption, and national security – namely the continued threat posed by Boko Haram. An article published by Freedom House[4] suggests that both candidates have probably been engaged in corruption in the past, stating that "the Economic and Financial Crimes Commission (EFCC), the country's main anticorruption agency, ha[d] initiated a number of high-profile investigations during Buhari's term", while "corruption by Buhari's allies in the APC often goes unpunished". The author also observes that "Atiku's long career has been marked by corruption allegations." Moreover, there is growing concern that military brutality is not being controlled by the president, after an incident in October 2018 when soldiers responded to rock-throwing protesters by shooting into the group and killing at least 45 people. This was just one incident against a backdrop of continuing violence against Muslim Shiites who are seen as a threat by the president who belongs to the Muslim Sunni. As a reaction to public outrage at the shooting, the military posted a video on Twitter showing US President Trump arguing that stones thrown at the military should be considered firearms (Figure 21).

[4] https://freedomhouse.org/blog/nigeria-heads-elections-february-governance-real-challenge

Local journalists have flagged the intense use of misinformation during the election season, saying fake news is currently "on steroids". One of the top stories identified as fake is an ongoing rumour that current president Buhari had died and been replaced with a clone. In particular, fake news has been used by most parties to attack their opponents. For example, Lauretta Onochie, a personal assistant for social media to the president, repeatedly shared a story on social media accusing the PDP of keeping Nigerians in poverty and then giving them food and money at election rallies (Figure 22). The picture she posted alongside these accusations turned out to be of an unrelated charity event, and her accusations were picked up by international news agencies. In lesser known instances, she reportedly accused Atiku of "shopping" for terrorists in the Middle East[5] and claimed he was on the watchlist of security operatives in the United Arab Emirates.[6] In addition, recently BBC reported on a video widely shared on Facebook accusing presidential candidate Abubakar of brokering a deal with Boko Haram in exchange for land and oil (Figure 23). Meanwhile, local news has reported on specific trolling efforts whereby young people are recruited by campaigns to set up news accounts every day which spread information/test certain tactics and engage in other trolling activity. While there are no official numbers in relation to the presidential campaigns, candidates running for governor or senator pay as much as 60,000 Naira (US$165) a month to people to handle their social media campaigns. Usually, small teams of about 12 people can run over 600 Twitter accounts.[7] Such articles do remain scarce, thus details of how and whether campaigns engage in such activities remain somewhat unclear. In reaction to these massive amounts of fake news, both social media platforms and journalists are looking for ways to combat the problem at hand. Fifteen media houses have founded an initiative called CrossCheck Nigeria, and work with the International Centre of Investigative Reporting to identify posts and claims as fake and publish the truth. However, participating journalists have admitted that their efforts represent no more than just a drop in the ocean, although they are not giving up. Facebook has announced that it would not allow foreign political ads related to the Nigerian Election. Facebook is working with the Nigerian Electoral Commission in an effort to minimize the spread of false information and fake accounts. While both Facebook and Twitter are public – which makes content control easier – another heavily used platform is WhatsApp, through which fake stories can diffuse covertly and unhindered due to its end-to-end encryption. The company has said it is building a "sophisticated machine learning system" to combat abuse and suspicious accounts, but it remains unclear how successful this system will be.

In December 2017, a broadly worded hate speech bill called the Digital Rights and Freedom Bill passed through the House of Representatives. Newspapers such as the local *The Guardian Nigeria* celebrated the passing of the bill which had been in the making since 2014 and is the first of its kind in Africa to specifically protect data privacy, free speech, press freedom and outline lawful interception and surveillance. The bill seeks to establish an independent National Commission for Hate Speech, but also stipulates death by hanging for those found guilty of any form of hate speech that has led to the death of another person.

[5] https://www.worldstagegroup.com/alleged-defamation-atiku-demands-public-apology-from-buharis-aide/
[6] https://tribuneonlineng.com/212888/
[7] https://www.newtimes.co.rw/africa/social-media-trolls-influencers-set-fight-nigerias-elections

In March 2018, when the bill passed the Senate, *The Guardian Nigeria* reported concerns by the International Press Centre that the bill could be a serious threat to the freedom of the press and safety of journalists. Similarly, the 2018 *Freedom on the Net* report raised concerns that the bill could be used by the government to silence online dissent. Thus, for now it remains unclear what effect the bill might have as, to date (20/02), it still awaits the president's signature.

As mentioned earlier, a major threat to Nigerian national security is the Islamist terrorist group Boko Haram. In early 2018, the UN reported that a total of 60,000 people had been displaced in the north-east of Nigeria due to ongoing hostilities. Reports, which are scarce, say the military continues to follow a strong and scientific psychological approach in their propaganda fight with Boko Haram, although they have still not publicly admitted to or explained any psychological operations they may be engaging with. The army's obsession with information control started back in 2013 when they offered battlefield tours to media outlets, and started pressuring newspapers on which stories they publish. According to the 2018 *Freedom in the World* profile on Nigeria by Freedom House, the military began monitoring social media in 2017 for hate speech and content undermining the government, military and national security. With a view to the upcoming 2019 election, they established a 'situation room' to monitor election violence. Together with the Cyber Warfare Command, activated on 4 February this year to disrupt terrorist propaganda, the 'situation room' is engaged with monitoring, identifying and countering various forms of fake news and propaganda disseminated by terrorists and other "subversive elements". Meanwhile, opposition parties have urged the army to stay away from the election after the president told them to be "ruthless" with those found interfering with the voting process.

Figure 21: Video published on Twitter reacting to military shooting in October 2018

Figure 22: Onochie accusing the opposition (the picture was later identified as belonging to an unrelated charity event)

Figure 23: Video posted on Facebook accusing the opposition of brokering a deal with Boko Haram

Source: BBC

PAKISTAN

Pakistan has witnessed an increase in the use of computational propaganda during the past year, particularly in connection with the general election on 25 July 2018 and the geopolitical tensions with India in February 2019.

The election was won by opposition party Pakistan Tehreek-i-Insaf (PTI) with Imran Khan becoming the new prime minister, taking over from the Pakistan Muslim League-Nawaz (PML-N). Pakistan was one of the countries mentioned by Mark Zuckerberg when, in 2018, he declared concerns about the risk of Facebook being used to manipulate elections; he announced that the platform would require all political advertisements to clearly mention who is paying for the message and for their identity to be verified (Mirbahar & Serrato,

2018). In the build-up to the election, news sources reported that all leading political parties were asking their social media teams to create fake profiles as part of their social media strategy (Sohail, 2019). According to *The Diplomat*, an anonymous social media executive of the PML-N reported that almost "everyone is running fake Facebook accounts and Twitter bots" to keep "pace with what others are doing" (Sohail, 2018). Social media managers from the PML-N, 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).

Parties across the political spectrum used bots and fake accounts to amplify messages ahead of the vote, according to research by the Atlantic Council's Digital Forensic Research Lab (2018). One PML-N candidate, Chaudry Riaz-ul-Haq, running for the NA-142 constituency, used the hashtag #NA142RIAZKA on Twitter, which gathered 3,144 mentions in 2 days. These mentions originated from 22 accounts, 17 of which were created in June 2018, with an average of 142 mentions per account, strongly suggesting automated activity. Similarly, the Pak Sarzameen Party's campaign #VoteForDolphin reached 11.2 million users, but with an average of 21 posts per user, also signalling potential bot activity. Alongside bots, fake accounts emerged such as @PakistansPoll, which claimed to be the official poll of Pakistan Twitter. From 2015 until June 2018, approximately 30% of the retweets it posted were from the PTI and its leader, Imran Khan. Democracy Reporting International found that, based on a sample of accounts, 52% of #PMLN accounts and 46% of #PTI accounts were likely bots. Digital Rights Monitor Pakistan (2018) monitored trending hashtags during the election period and found that, of the 800,000 tweets, retweets and replies for the 37 tracked hashtags, "almost all of them had high human-bot activity" and that "some human-bot accounts were also found directly engaged in incitement to violence against political rivals". Asad Baig, Executive Director of Media Matters for Democracy, noted that "bot platforms" were created to automate contribution to a hashtag for higher activity; as platforms such as TweetDeck were being used to automatically send a high volume of tweets.

Social media platforms continue to clamp down on this behaviour. Facebook removed 103 pages, groups and accounts for engaging in coordinated inauthentic behaviour on Facebook and Instagram as part of a network that originated in Pakistan in April 2019. These pages posted about the military, Pakistani interests, Kashmir communities, and current affairs. They also engaged in political news such as topics related to the Indian government, leaders and military. Facebook linked these accounts to employees of the Inter-Service Public Relations (ISPR) unit of the Pakistani military. The report alleged spending of US$1,100 on ads from May 2015 until December 2018, with 2.8 million accounts following one or more of these pages (Facebook, 2019).

There was a surge in computational propaganda in this reporting period following the heightened geopolitical tensions with India. On 14 February 2019, there was a suicide attack killing 40 Indian paramilitary police in Indian-controlled Kashmir. Responsibility for this attack was claimed by Pakistan-based terrorist group Jaish-e-Mohammed and resulted in further escalations from India (which claimed to have attacked terrorist camps in Pakistan) and Pakistan (which shot down an Indian fighter plane and captured its pilot – who was

eventually released). Disinformation proliferated on both sides, with false and doctored images spreading on Facebook and WhatsApp. In one instance, Pakistan's media outlets showed a 2016 image of a crashed Indian jet, claiming it was taken from that day (LA Times, 2019).

A common strategy in Pakistan is the use of disinformation to discredit political actors online, often by accusing them of blasphemy, a criminal offence which carries the death penalty (Freedom House, 2017). Fake profiles are problematic in Pakistan, particularly in the context of blasphemy accusations. *The Diplomat* (2018) reported the case of a journalism student named Mashal Khan, who was lynched in April 2017 by a mob that suspected he had uploaded blasphemous content on Facebook. Disinformation is also targeting political and human rights' activists, including rumours about activists committing blasphemy or working on "foreign" agendas (Shah, 2018). One such example of disinformation was the reporting that Cambridge Analytica had been hired by former prime minister Nawaz Sharif. This gained significant traction on social media, and even filtered into mainstream media and TV talk shows, with Bilawal Bhutto, Co-Chairman of the PPP, demanding the ruling PML-N should come clean over the allegations. This was debunked, and the initial reporting by Eurasia Future has since been retracted after being widely discredited (Ali, 2018).

Fake profiles are also used for military purposes. In January 2019, it was reported that a fake Facebook profile of an Indian army medic by the name of Anika Chopra had 'catfished' (when someone is lured into a relationship by a fictional online persona) an Indian soldier located on the India–Pakistan border (CNN, 2019). CNN reported that the Indian soldier had disclosed sensitive information such as troop and tank movements. The Indian army chief Bipin Rawat had said they were battling a widespread outbreak of Pakistani catfishing. It was further reported that up to 50 Indian soldiers had been catfished by this same account, which is believed to be operated by a member of Pakistan's Inter-Services Intelligence (Dhoundial, 2019). India's Union Minister Jitendra Singh declared in May 2018 that "studios" had 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).

Online surveillance is becoming increasingly sophisticated in Pakistan. Cybersecurity firm Lookout identified custom Android and iOS surveillanceware named Stealth Mango and Tangelo that were being utilized by a highly targeted intelligence-gathering campaign, which they believe to be operated by members of the Pakistani military (Lookout, 2018). It is claimed it is being actively developed, with the latest release as recent as April 2018. Similarly, Amnesty International uncovered an extensive network of fake social media profiles that were being used to infiltrate civil society organizations. For example, a prolonged campaign was launched against Diep Saeeda, a human rights' defender. These networks were used to infiltrate activist communities, luring them into giving away their Facebook or Google log-in credentials, or to download malicious spyware (Amnesty International, 2018).

Pakistani authorities curb political discourse using censorship and blocking in order 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 sexual content and nudity, and tools used to circumvent censorship or protect privacy (Freedom House, 2017). Social media content has been restricted during religious and national holidays; and, in November 2017, social media platforms were suspended nationwide for 2 days in the wake of protests that turned violent (Freedom House, 2018). In April 2018, news website NayaDaur was blocked for over a week before the Pakistan Telecommunication Authority unblocked it. No reason for the blocking was given, but it followed the publication of an article sympathetic to the Pashtun human rights' movement (Freedom House, 2018). In June 2018, during the run-up to the general election, individuals trying to access a website operated by the Awami Workers Party were told the website was not accessible as it "contains content that is prohibited for viewership from within Pakistan" (Jahangir, 2018). According to official figures, there are at least 831,002 sites blocked in Pakistan; with 769,947 over pornographic content and 34,762 over blasphemy (Jahangir, 2018).

New cyber agencies have been created in the last year within the Pakistani government. In early 2018, US$196,000 was allocated to establish a Cyber Patrolling Unit to be run by the Federal Investigation Agency (FIA), with the aim of targeting online child pornography (Freedom House, 2018). Under Pakistan's cybercrime law, the FIA is designated with monitoring hate speech. In February 2019, Information Minister Fawad Chaudhry said that the government has "prepared a mechanism" to control hate speech on social media. They would also introduce a new authority, the Pakistan Media Regulatory Authority, to enforce regulations for digital, print and electronic media (Geo News, 2019).

QATAR

Qatar is consistently ranked not free in Freedom House's annual rankings of Internet freedoms, as the hereditary Emir, Sheikh Tamim bin Hamad Al Thani, holds all executive, legislative and judicial authority. Four fifths of the population are non-citizens with no political rights, few civil liberties and limited access to economic opportunities (Freedom House, 2018). Efforts at computational propaganda from Qatar are poorly documented, aside from a small number of media reports. Qatar has been at the centre of ongoing diplomatic tensions with other states in the region, which have often spilled into the online domain.

Qatar faced a diplomatic crisis on 23 May 2017, after the state-run Qatar News Agency (QNA) posted controversial statements allegedly made by the Qatari Emir Sheikh Tamim Bin Hamad Al Thani. The comments affirmed good relations with Iran, the Muslim Brotherhood, Hezbollah and Hamas. Qatari officials quickly denied that Thani had made these comments, claiming that QNA and associated social media accounts had been hacked. Nonetheless, Saudi Arabia and the UAE media dismissed the hacking story and accused Qatar of supporting terrorism (Washington Post, 2017). This caused Gulf Cooperation Council (GCC) members to sever diplomatic ties with Qatar, which coincided with an online information battle. An added layer of complexity to this issue came when US investigators

announced that it was Russian hackers that had breached QNA and planted the false news report that triggered the crisis, allegedly orchestrated by the UAE (CNN, 2017).

Qatar has been the target of ongoing computational propaganda and disinformation campaigns (see this report's profiles on Saudi Arabia and the UAE). News outlets have reported that the rift between the Saudi government and Qatar is "frequently bot-ridden" (NBC, 2018) and that 17% of a random sample of Arabic tweets mentioning Qatar in a sample in 2017 were sent by automated accounts (Washington Post, 2018). In August 2017, Saudi Arabia's 'king of disinformation' Saud al-Qahtani said that the hashtag #LeaveTamim was trending in Qatar, reflecting how Qataris wanted to oust their ruler, however this hashtag was mostly generated by anti-Qatar bots, signalling that Qatari Twitter trends are subject to international manipulation (Jones and Abrahams, 2018). Reuters reported that "online attacks against the small Gulf state surged" following the diplomatic boycott of Qatar in June 2017 (Reuters, 2017).

Despite these attacks on Qatar, there is evidence that Qatari computational propaganda added to this dispute too. In 'The Online War between Qatar and Saudi Arabia', the BBC reported that the hashtags 'Tamim the Glorious' and 'Qatar is Not Alone' appeared on Twitter's trending homepage, and that "the majority of tweets using these hashtags were pushed by fake accounts". On #Tamim_the_Glorious, one account (@sabaqksa) had 201 retweets in the space of a couple of seconds, which Ben Nimmo, from the Atlantic Council's Digital Forensic Lab, said was "not a normal pattern of behavior" (BBC, 2018). This hashtag was featured in 90,000 retweets in the space of a few hours (Nimmo, 2018). Another surge of traffic on this hashtag saw one hundred accounts posting 1,410 times in a five-hour period, which Nimmo called "utterly implausible" that humans could have operated. In the Columbia *Journal of International Affairs*, Nimmo states that Twitter bots were "deployed to boost messaging on both sides of the diplomatic dispute between Saudi Arabia and Qatar", and that some of these "appeared commercial" botnets whereas others were "locally focused". On the hashtag 'Qatar is not alone' (#قطرليستوحدها) a scan of tweets highlights two significant spikes in activity in the initial hours of 24 May 2017, following the diplomatic incident. Traffic more than doubled in the course of a single minute, suggesting possible botnet involvement (Nimmo, 2018). There also appeared to have been foreign involvement, as a network of bots "whose primary language and focus appears as Turkish" drove spikes in hashtags. Nimmo concluded that the focus of these Arabic-language hashtags was clearly local and regional rather than international; an attempt at messaging to the domestic population rather than the non-Arab world (Nimmo, 2018). The online information battle against Qatar in the Gulf is also explored in a paper by Marc Owen Jones (2019).

It is worth noting that Bahrain has publicly accused Qatar of computational propaganda, although the following reports must not be taken at face value due to the nature of the ongoing diplomatic dispute and authenticity of the sources. The Kingdom of Bahrain Ministry of Interior released a statement on 21 July 2018 which claimed Bahrain was subject to systematic targeting to "compromise its national interests by influencing public opinion" and hitting the economy by "circulating false information through fake accounts run from Qatar" (Bahrain Ministry of Interior, 2018). The *News of Bahrain* was more specific and claimed that "thousands of fake social media accounts are being used by the Qatari regime

every day to defame Bahrain" (News of Bahrain, 2018). While these claims provide no evidence, combination with more verifiable reports that Qatar is engaging in a degree of computational propaganda suggests that Qatar has developed its social media capabilities over the past few years.

RUSSIA

Reports of Russian manipulation of social media have dominated the international news cycle for the past few years. Russia is often cited as one of the first and most sophisticated actors to engage in computational propaganda. These policies need to be understood both in the context of the Russian presidency's consolidation of power and the broader historical background. President Vladimir Putin has been the dominant figure in Russia's political landscape since his election as president in 2000 and he was re-elected for a third term in March 2018 (BBC, 2018). Throughout his presidency, Putin has restricted the independence of various state institutions and the media, which has been accompanied by increasing nationalism and hostility to the West.

Computational propaganda efforts are closely intertwined with traditional media outlets. Russian media is dominated by channels that are either run directly by the state or owned by companies closely linked to the Kremlin, such as Russia Today (rebranded as 'RT') and Sputnik (BBC, 2018; DFRLab, 2018). Within the Russian government narrative, 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 designate RT and Sputnik as 'foreign agents' 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; DFRLab, 2018). RT's parent company, TV-Novosti, is registered as a state-owned Autonomous Non-commercial Organization with the Russian Ministry of Justice and is almost entirely funded by the state budget (99.5%–99.9%) (DFRLab, 2018). RT's editor-in-chief has even gone as far as describing RT as an "information weapon" used in "critical moments" (DFRLab, 2018). To this end, according to Ben Nimmo, RT "subordinates journalism to one-sided reporting and selective interviewing to support the Russian government's narratives and 'conduct the information war'" (DFRLab, 2018).

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 cybersecurity which encompasses the dissemination of information "harmful to the spiritual, moral and cultural spheres of other states". (Gjelten, 2010; Franke, 2015). Russia's approach to the information sphere encompasses both offensive military-cyber capabilities and information, and content strategies – an approach that distinguishes Russia's IW (information warfare) tactics from the West – and considers itself to be engaged in 'full-scale information warfare' (Giles, 2016).

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 techniques. Many other countries, both targets and allies, have consciously or unconsciously begun to imitate these techniques (Diresta, 2018). Media often cite that Russia's 'playbook' has been globally adopted (Frenkel, 2019). These tactics often follow the pattern of "dismiss, distort, distract, dismay" (DFR Lab, 2018). This involves sophisticated trolling on news sites, fake hashtags and online campaigns, and close coordination with other media operations (Helmus et al. 2018). The result is a blend of attributed and non-attributed elements, allowing a scale of complexity and plausible deniability. Furthermore, while media often focus on Russia's direct electoral interference, such as the US presidential election in 2016, Russia's information warfare has also sown long-term divisions in foreign societies. For example, in May 2016 a handful of protesters opposed the opening of a library at an Islamic Centre in Houston, Texas; the rally – 'Stop Islamization of Texas' – had been promoted by a Facebook page called 'Heart of Texas', operated from the Internet Research Agency (IRA) in St Petersburg (CNN, 2017).

The IRA has been one of the principal players in Russian efforts to fight the information war. Founded in 2013, it hired 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) (See this report's profile on Ukraine.) 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 US$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).

In February 2018, the US special counsel investigation into Russia's interference in the 2016 US election, led by Robert Mueller, indicted 13 Russian nationals and three organizations for "conspiracy" to illegally influence the US 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). The 2018 indictment offers the best insight into the organizational capacity of the operations. It is claimed the IRA operated with a monthly budget of as much as US$1.25 million and spent thousands of dollars a month buying political advertising (BBC, 2018). During the election, more than 99% of all engagement came from just 20 Facebook pages controlled by the IRA – including 'Being Patriotic', 'Heart of Texas', 'Blacktivist' and 'Army of Jesus' (Washington Post, 2018).

By 2014, the IRA expanded their activities to target the US 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 travelled to the US "to collect intelligence for their interference operations" according to the indictment. The IRA 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 DC, 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."

Alongside the IRA, other Russian state organizations have assisted in computational propaganda efforts. Mueller's indictment also alleged that the GRU, Russia's military intelligence, hacked emails from Democratic Party staff and allies, before publishing them online and promoting them through a network of fictitious social media personalities (DFRLab, 2018). Ben Nimmo describes how this operation combined hackers with a "fake grassroots campaign, fake social media accounts, non-existent journalists, and a dedicated website" (DFRLab, 2018). This operation was smaller than that of the IRA, focusing mainly on mobilizing African American opinions (utilizing '#BlacksAgainstHillary'), and supporting Russian military operations in Syria – as well as working closely with hacking units (DFRLab, 2018). The role of the GRU was further brought to light in October 2018 following the arrest of four hackers from the GRU for targeting the World Anti-Doping Agency (WADA), providing leaks which were amplified through government outlets, Russian media and Internet trolls (DFRLab, 2018).

Reports presented to the Senate Intelligence Committee in December 2018 have shone a light on the reach of Russia's interference in the 2016 presidential election. New Knowledge determined that the IRA reached 126 million on Facebook, 20 million on Instagram, 1.4 million on Twitter, and uploaded 1,000 videos to YouTube (New Knowledge, 2018). The Computational Propaganda Project found that Russian interference continued after the election ended, that targeting began as early as 2012, and that the IRA sought to divide American voters along lines such as race, ethnicity and identity (Howard et al. 2018).

Despite the Mueller indictment, Russian accounts have been accused of continuing to interfere with a variety of US 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 US context, even if these tend to draw the most attention – Russian efforts now have a global reach. Propaganda campaigns have been

conducted in Russian, English, Arabic, French, Czech, Georgian and many other languages (Helmus et al. 2018). Snegovaya (2017) details a sophisticated campaign to target the Russian minority in the 2017 German elections. A *Guardian* investigation found that in the UK, tweets from members of the IRA "troll army" were quoted more than 80 times across British-read media outlets (Hern, Duncan, and Bengtsson, 2017). In April 2017, Russian bot activity increased following the chemical attack in Khan Sheikhoun, Syria, which UN investigators concluded was perpetrated by Russia's ally, the Syrian government. The hashtag #SyriaHoax was the number one trending topic on Twitter, boosted by an army of inauthentic accounts (The Syria Campaign, 2017) (See this report's profile on Syria). In March 2018, the poisoning of Sergei Skripal, a former Russian intelligence operative, in Salisbury, UK, was followed by a wave of online propaganda and disinformation. UK government analysis claimed that they uncovered a 4,000% increase in the spread of propaganda from Russia-based accounts since the attack, many of which were identified as automated bots (Guardian, 2018). Official Twitter accounts were even involved – with the Russian Foreign Ministry's official account mocking the British government with tweets accusing them of blaming Russia for everything, even the weather (DFR Lab, 2018). In December 2018, Russian accounts were implicated in using the hashtag #giletsjaunes, the French name for the Yellow Vest protest movement (Bloomberg, 2018). The DFR Lab had previously uncovered extensive Russian information operations in France related to President Emmanuel Macron, origination attributed to the IRA (DFR Lab, 2018). In all, Russian information operations have been uncovered in some form in a large number of European states (Kremlin Watch, 2018; Dorell, 2017), Canada and the USA (Time, 2018; Dorell, 2017), Australia (ABC, 2018), and many other countries are beginning to investigate the extent of Russian interference.

Propaganda has been discovered on almost every major media, social media and technology platform. Ukrainian soldiers have been targeted with propaganda by SMS text messages on their mobile phones (DFR Lab, 2017). NATO's Strategic Communications Centre uncovered extensive inauthentic activity on Russian service VK (VKontakte) among Russian-speaking populations in Estonia, Latvia, Lithuania and Poland (NATO StratCom, 2019). Reports indicate that alongside the major social media platforms, Instagram, Vine, Pinterest, SoundCloud, Pokémon Go, Tumblr, Reddit, Google (Google+, Gmail, Voice, Ads), Meetup, Medium, Gab and PayPal were all used to some extent in the 2016 presidential election operation (Business Insider, 2018; New Knowledge, 2018).

Social media platforms, in the face of increased public pressure, are beginning to take measures to counter online information operations. In October 2018, Twitter released the account and related content associated with Russian information operations since 2016, including 3,841 accounts affiliated with the IRA and 9 million associated tweets (Twitter, 2018). Facebook removed 1,907 accounts linked to Russia in March 2019 (Reuters, 2019; Facebook, 2019). In May 2019, Facebook removed 97 Facebook accounts, Pages and Groups for targeting Ukraine, and 21 accounts, Pages and Instagram accounts for targeting Austria, the Baltics, Germany, Spain, Ukraine and the UK – campaigns run on these pages and accounts were suspected to be a Russian intelligence operation (Facebook, 2019).
While there is limited evidence, it is likely that Russia is investing in research and development in information operations given its central role in domestic and foreign policy.

One particularly concerning development is the use of Artificial Intelligence (AI). Alina Polyakova wrote that in the near term, Russia will "ramp up the development of AI-enabled information warfare" (Brookings, 2018). It is claimed that Russia's social media campaign to influence elections employed "relatively basic AI" but also "machine learning… to deploy armies of bots, targeted ads" (Wilson Quarterly, 2018).

SAUDI ARABIA

In the past year, the Kingdom of Saudi Arabia has seen a decline in Internet freedom and an increase in computational propaganda efforts. The kingdom is connected to the Internet through two country-level data service providers, and because all user requests that arrive through Saudi Internet Service Providers (ISPs) travel via these servers, they are subject to censorship at this centralized point. Websites that are judged to be harmful, illegal, anti-Islamic or offensive are routinely blocked (Freedom House, 2018).

Computational propaganda, most notably through Twitter bot networks and disinformation, has increased – specifically surrounding the murder of prominent Saudi critic Jamal Khashoggi in October 2018. The kingdom is ruled by Crown Prince Mohammed bin Salman (MbS), and at the centre of the kingdom's computational propaganda is his close adviser, Saud al-Qahtani. The government is suspected of employing an 'electronic army' that posts pro-government messages, inflames sectarian tensions, targets foreign states such as Qatar, and trolls Saudi critics.

Following an economic boycott against Qatar in June 2017, al-Qahtani encouraged Twitter users to use the hashtag #TheBlackList (#السوداء_القائمة) to turn in fellow citizens who sympathized with Qatar and vowed to follow every name that was reported to him. Saudi writer Turki al-Ruqi accused al-Qahtani of acting like an Internet troll when he launched social media campaigns to intimidate dissidents (Alaraby, 2018). Bot networks changed their location to Qatar and propelled anti-government hashtags to the top of Qatari Twitter trends, which led to foreign Twitter users concluding Qataris were demanding a change of leadership (Al Jazeera, 2018). Research at Columbia University discovered that, on both sides of the row, automated networks of high-volume Twitter accounts amplified their messages and boosted their hashtags. Researchers' analysis signalled that some of these botnets may have been commercial and hired from abroad, whereas others were made to appear as locally based in both Saudi Arabia and Qatar (Ben Nimmo, 2018).

There is undoubtedly a large presence of automated bots on Twitter in Saudi Arabia. Marc Owen Jones, an Assistant Professor in Middle East Studies and Digital Humanities at Hamad bin Khalifa University, Doha, who investigates propaganda on Twitter, suggested that half the active Twitter users in the kingdom may be bots (Al Jazeera, 2018). His analysis has found that bots in the kingdom focus on hashtags surrounding domestic issues such as #Saudi, #Riyadh and #AlQatif, often in support of Saudi government or foreign policy, but also international targets such as the #Bahrain and #Yemen hashtag, or in propagating sectarian rhetoric (Marc Owen Jones, 2018). He alleges that the scale of this operation is enormous, with dormant Twitter accounts used as 'fake followers', including "potentially up to a million of these accounts". According to his data, 70–80% of Arabic-language tweets

containing the word 'Saudi' in the past four months were posted by bots (Foreign Policy, 2018). There were even reports of the Twitter accounts of deceased celebrities being used to spread pro-Saudi propaganda. The Twitter account of David Schwartz, a weather channel meteorologist who died in 2016, had his Twitter handle used to post pro-Saudi messages – possibly because the Twitter account was verified yet no longer being used (Alaraby, 2019). Further, there is evidence of the commercialization of these bot networks: a BBC investigation found that companies in Saudi Arabia were offering to artificially boost the popularity of hashtags on Twitter, quoting the equivalent of £150 to make a hashtag trend for several hours (BBC, 2018).

The New York Times (2018) reported that Saudi Arabia had been attempting to infiltrate Twitter itself. At the end of 2015, Western intelligence notified Twitter executives that the Saudis were grooming an employee, Ali Alzabarah, to spy on the accounts of dissidents.
The murder of Jamal Khashoggi is fundamentally linked to the ongoing computational propaganda efforts in the kingdom. Khashoggi was a prominent Saudi dissident and had been writing articles critical of Saudi Arabia for *The Washington Post*. On 2 October, he was murdered by Saudi intelligence officials in Riyadh's consulate in Istanbul, Turkey. A United Nations human rights investigation into the murder found evidence that it was a premeditated killing, planned and perpetrated by officials of Saudi Arabia (Al Jazeera, 2018). There is even evidence of al- Qahtani's involvement in the Khashoggi murder, as Reuters reported that, according to a high-ranking Arab source, al-Qahtani was video-called via Skype in the room in the Saudi consulate prior to Khashoggi's murder (Reuters, 2018). Khashoggi had been attempting to combat online abuse, and had wired US$5,000 to Omar Abdulaziz, a Saudi dissident in Canada who was creating a volunteer army known as the 'Electronic Bees' to combat the government Twitter trolls (New York Times, 2018). *The Independent* reported that Khashoggi was at the heart of an 'online army' of Saudi activists fighting a misinformation cyberwar.

Following Khashoggi's murder, automated bots flooded social media, intending to cast doubt on any allegations that the kingdom was involved in his death (Reuters, 2018). At least 53 websites, including alawatanews.com, were part of a network that posed as authentic Arabic-language media outlets to spread disinformation about the Saudi involvement in Khashoggi's murder (Reuters, 2018). The hashtag announcing Khashoggi's 'kidnapping' disappeared from the list of top trends in Saudi Arabia after a few hours, suggesting an army of accounts had worked to deliberately bury it (Independent, 2018). The hashtag #UnfollowEnemiesOfTheNation was mentioned 103,000 times in the days following the murder (Washington Post, 2018), and analysis determined that there were hundreds of postings per second (BBC, 2018). Ben Nimmo, from the Digital Forensic Research Lab, found that 96.3% of the uses of these hashtags were retweets, suggesting a coordinated effort from bots or a retweet farm. On 14 October, Arabic hashtags topped the global trends; 'we all have trust in Mohammed bin Salman' was featured in 250,000 tweets, and 'we have to stand by our leader' in 60,000 tweets. Some of these networks have existed and been spreading propaganda since 2012, whereas others appear commercial and were specifically rented for the occasion (CNN, 2018).

In response to this wave of automation, NBC News identified a large number of Twitter accounts created in quick succession in November 2016 and being utilized to spread pro-Saudi messages (NBC News, 2018). In response, hundreds of these pro-Saudi accounts were deleted by Twitter in October 2018 (CNN, 2018). It claimed that these pro-government bots had been part of an online propaganda campaign since 2016 (Independent, 2018).

A bogus Saudi fact-checking account called 'Middle East Guardians' published a photo that it claimed had been altered to include Khashoggi's Turkish fiancée, suggesting that she had not been present before his murder and has in fact no relationship to Khashoggi at all. She has been a prominent voice accusing Saudi Arabia of being involved in Khashoggi's death and the claim of 'Middle East Guardians' has led to her becoming a target for people who claimed that the whole incident was simply a set up to make Saudi Arabia look bad. However the photo was disproved, and the Twitter account suspended. The extent of disinformation even led to Reuters falling for a fake news article about the firing of a Saudi general consul, and they had to retract their article (Poynter, 2018).

The hub of these operations is reportedly the Center for Studies and Media Affairs in Riyadh, run by Saud al-Qahtani (Washington Post, 2018). It is alleged that, as early as the 2000s, al-Qahtani was tasked with building an 'electronic media army,' a network of surveillance and social media manipulation to advance the crown prince's agenda and suppress his critics, and which was further fuelled by the Arab Spring. He is reported to have developed an "army of flies" and was labelled by activists as the 'troll master', 'Saudi Arabia's Steve Bannon,' and 'minister of disinformation' (Washington Post, 2018). This trolling is not confined to simply bot accounts: Ghada Oueiss, an Al Jazeera journalist, has been called a "prostitute" by an account with 338,000 followers – verified as Saudi Minister Abdullatif al-Shaikh (Washington Post, 2018).

SERBIA

Most efforts around computational propaganda in Serbia seem to originate from the ruling Serbia Progressive Party (SPP) and its leader Aleksandar 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.

Government control over social media is tight: at the time of the 2014 floods, when the government response was absent or uncoordinated, citizens used social media to organize aid and volunteers. Several, however, were called in by the police for questioning and threatened with charges such as "spreading panic" for their posts. Meanwhile, government tabloids spread fake clickbait news about hundreds of dead bodies floating on rivers or Roma gangs on the loose. Generally, these state-sponsored or party-owned tabloids are responsible for spreading right-wing discourse throughout the country, which facilitate violence and hatred. Most domestic news sources are dominated by black and white stories,

which either praise Russia, and condemn the West, while outlets with a more pro-West lens have virtually no impact in Serbia. The country generally has an underdeveloped media landscape with relatively little unbiased and accurate news coverage: most outlets simply copy and paste government statements, and reports by Kremlin-backed media, such as Sputnik, are widely quoted by Serbian mainstream media. In addition, a network of online influencers and outlets shares Sputnik's content affording the platform a broad reach across the country. As a reaction to increasing disinformation instigated by foreign parties such as Russia, in 2018 the BBC launched its news service *BBC Serbia*.

Ever since Vucic came to power in 2014, freedom of both the press and journalists in Serbia has continued to decline. Before becoming prime minister, Vucic was Minister of Information from 1998 to 2000 and as such the main force behind a law on information which fined journalists criticizing the government and banned foreign TV networks. Ironically, Vucic took part in the Davos panel 'Media Freedom in Crisis' in late 2018, which shocked many in the press and journalistic community. On the panel, Vucic talked about the critical situation of press freedom in his country and said he was working to solve the problem, even though matters have actually worsened since his ascent to power. Local journalists in particular feel this event has led to their struggle being ignored internationally. Retracting freedom of information in general seems to be a trend across the wider Balkan region; Croatia and Montenegro have also taken steps to reduce the freedom of information and increase the likelihood of state secrecy: 'Right to know' legislation, which was introduced in several Balkan countries in the mid-2000s, is increasingly infringed or simply ignored.

As of March 2019, Serbians are becoming more vocal in their resistance against the state's control of the media and elections, marching in daily "anti-dictator" protests demanding the president's resignation. They are also demonstrating for more media freedom, more coverage of opposition groups by public broadcasters, and an end to attacks on journalists and opposition figures. These protests started after opposition leaders were beaten and public broadcasters refused to cover subsequent demonstrations condemning the violence. However, next to this progressive force there are also right-wing extremist groups who participate in these protests. The civic democratic section of the protests is at the very least tolerating the extremists, so it remains to be seen whether these new movements are really about democratizing the country and media, or whether it is new interest groups who mainly want to control the media themselves.

Meanwhile the Serbian government is focusing its efforts on increasing military capability and knowledge. Under the 2019 budget, it allocated US$907 million (1.75% of its GDP) to the military and additionally received fighter planes donated by Russia and Belarus. Moreover, defence and security lessons were re-introduced to school curricula. Serbia did renew its pledge to hold a neutral position in early 2019 and not side with Russia or the NATO. It remains unclear, though, whether or how Serbia may have increased its military intelligence capability through the application of the increased budget.

SOUTH AFRICA

Generally, South Africa is considered a free country with a free online space which has established itself as a platform for political mobilization and debate. Nonetheless, the country has had its fair share of incidents relating to cyber troops and misinformation. According to a survey from the Edelman Trust Barometer (reported by News 24), 62% of respondents said they are unable to separate fake news from real news and 54% reported that it is increasingly more difficult to know whether news is credible or not. Additionally, only 34% trust in media in general, with 69% of South Africans being worried that fake information or news will be used as a weapon in the future. These numbers, in light of the national election on 8 May 2019, are cause for concern. In terms of Internet penetration, nearly two thirds of South Africans are able to access the Internet according to the 2018 *Freedom on the Net* report. However, there is a great urban–rural divide; for example, 71% of people have Internet access in Western Cape, but only 44% in the province of Limpopo. The majority of Internet users go online through their mobile phones and one of the main obstacles to Internet access is high costs. The government has been trying to establish free Wi-Fi access but, so far, their efforts have only been successful in metropolitan areas. Additionally, 9 of the 11 official languages of the country are underrepresented online.

The South African government is generally not involved in controlling access or censoring content online. The five major undersea cables connecting the country to international Internet are all operated by private companies. In recent years, however, several governmental officials have announced the intention to regulate social media, on the pretext that it is increasingly used to spread false information. In 2017 the Minister of State Security acknowledged that such regulation could be seen as interference with human rights online. So far, there are is no specific legislation for social media regulation. Nevertheless, the government voted against a UN resolution for "the Promotion, Protection and Enjoyment of Human Rights on the Internet" in 2016, siding with countries such as China and Russia. South Africa said they thought the resolution failed to account for hate speech and incitement, which are great challenges for the country's post-apartheid society.

There are two major bills which could affect governmental censorship and surveillance rights. The first one is the Film and Publications Amendment Bill, which passed the National Assembly in 2018 and again in 2019 but still needs to be signed by the president to be passed into law officially. The bill is supposed to update the definition of child pornography in an effort to protect children from harmful content, but it could be abused to censor content online by, for example, regulating platforms such as YouTube and blocking websites. The second is the Cybercrimes and Cyber Security Bill, which was first drafted in 2015 and quickly spiked controversy for its vague language. It includes passages which penalize the dissemination of "data messages which are harmful", which is defined as information that is "inherently false" without further specification. Critics see it as empowering the state's ability to surveil citizens. Pushed by the Department of Justice and Constitutional Development, the National Assembly passed an amended bill in 2018, which excluded several cybersecurity sections.

While these bills could increase surveillance, South African citizens have been empowered to legally protect their privacy as well: since 2013 the Protection of Personal Information Act allows individuals to bring civil claims against those who defy the act. Punishment

includes prison sentences and fines of up to $650.000 (ZAR10 million). In recent years there have been incidents of content control and removal, although not all of it is instigated by the government. For example, in July 2017 the website Black Opinion was taken down by one Internet service provider after it received complaints that the site fostered racism – although the site was then restored two weeks later.

This latter example seems to be related to one of the biggest events on social media amplification involving the billionaire Gupta family. Among the richest people in the country, the Gupta brothers have long been accused of corruption and influencing the administration of ex-President Zuma (he resigned in 2018). It was revealed that the Guptas paid the British PR agency Bell Pottinger to improve their media image amid a watchdog investigation into their alledged 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 US$2 million for the whole operation. The Bell Pottinger strategy sought to divert attention from the Guptas' influence on government, hiring a Florida-based digital marketing firm, Devumi, which deployed botnets to amplify the "white monopoly capital" phrase and target journalists who were critically reporting on the influence the Gupta brothers had on the Zuma administration. 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. Twitter has shut down around 900 bots that systematically spread white monopoly capital content and promoted fake-news websites in the form of botnets and sock puppets. However, it appears that new fake accounts are set up as quickly as they are taken down.
Official investigators found that the bots were not effective. On the contrary, accounts – whether automated or human-administered – which tweeted in favour 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. Nonetheless, organizations such as Freedom House announced concerns that such activities are increasing self-censorship among critical reporters in the country.

While these efforts were driven mostly by private individuals, there are examples of the government trying to control narratives directly, most prominently through the state-owned broadcaster SABC. Several board members of the company are part of the African National Congress (ANC), which is the governing party. According to some critical writers, this entanglement has led to the downfall of the broadcaster, calling their programme ANC propaganda and pointing out the dire financial situation the company is in. In light of the upcoming election, one can only expect that ANC-controlled programmes on SABC will increase among efforts to win the election. During 2016 local elections, the party allegedly spent US$2.75 million on a "black ops" room or "war room" (Comrie, 2017) to run misinformation campaigns against their opponents, with SABC one of the main channels to reach millions of rural voters. Additionally, the black ops included seemingly neutral news sites (e.g. The New South African, which no longer has a website) and chat shows, using

around 200 influencers on social media (mainly Facebook and Twitter), and producing fake opposition campaign posters. Reportedly, the 'room' never went into full operation due to mismanagement and lack of funding, and the ANC has since attempted to distance itself from the campaign. This includes threatening journalists who were investigating the story (Figure 24). Given the expected amount of misinformation and fake news that would arise in relation to the elections, Google started training political parties, journalists and editors to spot and fight fake news as part of a US$300 million international initiative announced in March last year.

Figure 24: Nikadimeng who owns one of the companies supposed to run The New South African messages to journalists

An apparently threatening text message to an amaBhungane reporter from Nkadimeng

Source: https://www.news24.com/SouthAfrica/News/exclusive-the-ancs-r50m-election-black-ops-20170124

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 whistleblower also said that he was paid US $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).

In October 2018 South Korea announced a crackdown on "Fake news". Police and prosecutors were encouraged to investigate individuals generating fake news with malicious intent. According to the New York Times, opposition lawmakers denounced this move as an attempt to silence criticism, especially on platforms such as YouTube where critics commonly post videos.

SPAIN

Spain is considered a free country and stable democracy because in general the government does not engage in any form of limiting Internet access, blocking or censoring. Internet penetration rates are high, with about 93% of the population using the Internet daily in 2017. The country has a free and independent media, although most outlets, radio or TV stations are part of bigger corporations, which has caused some concern as a potential threat to media freedom and impartiality. Similar concern was raised in 2015 when the government passed the Citizen or Public Safety Law which was dubbed the "gag law" by critics because it established large fines for offences such as publishing pictures which could threaten the safety of police officers or protected buildings (i.e. any pictures of police protecting such buildings). During the Catalan referendum in particular, journalists faced severe penalties under this law while reporting on police activity in Catalonia. According to research by Amnesty International published in late 2018, since the law came into effect the government has made almost €25 million and an average of 80 people are subject to fines each day. At the same time, the governing socialist party PSOE – with support from the left-wing party Podemos – announced plans to reform the law and loosen its grip and intrusive nature, but they reversed their plans at the last minute and significantly watered down their proposal.

The most prominent and profound incident which recently caused controversy in Spain was the 2017 Catalan independence referendum. The situation leading up to and following the vote became chaotic with disinformation and fake stories flooding the debate. The national public broadcaster RTVE was criticized by its own journalists and its news council (overseeing the broadcaster's impartiality) for biased coverage of the referendum. Generally, journalists were pressured by both sides to cover the events in a certain way. Additionally, Spanish authorities tried to block websites which shared information about the referendum. Observers of news and information at the time said they had never seen anything like it. It is most likely that the extreme number of fake stories and hate content was due to the sensitive and important nature of the referendum. Many pictures alleging police brutality were circulated at the time aimed at discouraging people from voting, even though most of these pictures were completely unconnected with the Catalan context (Figure 25). However, this is not to say that police brutality was not an issue. Moreover,

fake tweets from politicians were shared as well as stories claiming Spanish tanks had been deployed to Catalonia. The use of WhatsApp was also hugely problematic because fake stories circulated freely in private chats making it extremely difficult for fact-checkers to debunk false claims.

The Spanish central government accused Russia of meddling with the Catalan referendum through their own groups and news outlets such as Sputnik and Russia Today (RT). In November 2017, the government announced their intelligence indicated Russian-based groups used social media to spread misinformation: about 50% of the accounts came directly from Russia and 30% from Venezuela. Russia declared Spain's accusations part of the Western hysteria against Russia affecting national politics around the world, while Catalans mocked the central government saying they hardly needed Russian reminders of their grievances. Spain's Foreign Minister has stated that there is indeed no concrete evidence that Russia was involved. Nevertheless, the EU's fact- checking task force found evidence of Russian-backed media spreading disinformation about the Catalan situation (Figure 26). Half of the stories shared by RT a day before the referendum were about police violence with headlines such as "Powerful videos: the brutal police repression against voters in the Catalan Referendum". In addition, Researcher Javier Lesaca of the Washington State University analysed more than 5 million social media messages sent from 29 September to 5 October 2017 by Sputnik and RT and found an "entire army of zombie accounts that are perfectly coordinated". Lesaca too found that part of the networks used were previously employed in Venezuela (Figure 27). The Madrid-based think tank Royal Elcano Institute called these activities part of the Russian information "war" in Catalonia. Other experts such as Klaus-Jürgen Nagel from the University of Barcelona (Universitat Pompeu Fabra) said such a claim is an exaggeration and Russia simply provides information from their geopolitical perspective. Ben Nimmo of the American think tank Atlantic Council also thought it unlikely that the Russian-backed media outlets had any specific orders from the Kremlin.

In early 2019, the trial of 12 Catalan separatists began in Madrid. The government launched an EU-wide campaign to convince people that the trial is not politically charged. For example, the Spanish ambassador to London spoke out saying that the pro-independence regional government of Catalonia had launched a "massive campaign of disinformation" with "the principal underlying message [...] that Spain is not a democracy" to criticize and delegitimize the trial and the Spanish central government. According to the ambassador, Spain intended to fight back with transparency (they are livestreaming the trial for example) and its own campaign called 'This is the real Spain' to advertise the country's diversity in opinion and inclusiveness.

Spain is indeed experiencing a new-found diversity – at least in their political system. Traditionally, the country has consisted of a two-party democracy, however, the latest general election held in April 2019 was probably the ultimate indication that such times are in the past. The general elections were called by Prime Minister Pedro Sánchez after right-wing and Catalan separatist parties rejected the 2019 budget proposed by his minority coalition. Observers expected right-wing parties to significantly profit from the election and in fact the right extremist party Vox took seats for the first time, but the winner was

Sánchez' Socialist party (PSOE) which received 28.7% of the votes. They were followed by the conservative People's Party (16.7%) and the centre-right Citizens party (15.9%).

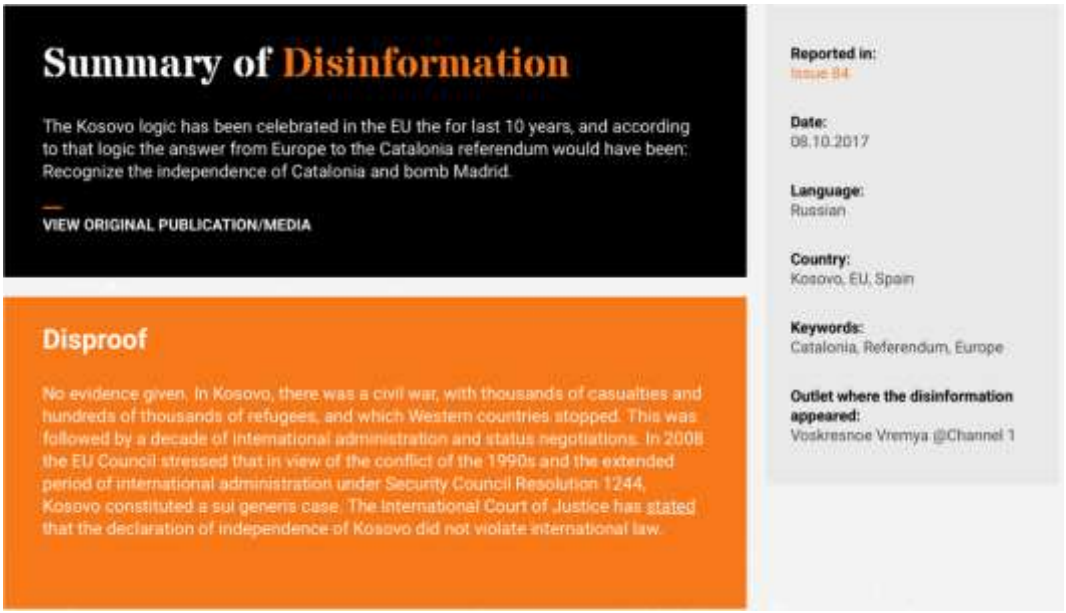
Days before the election, Facebook quietly took down 3 far-right networks for fake and duplicated accounts, which ran about 30 Facebook pages reaching over 1.7 million Spaniards. Amongst the content shared were fake stories and doctored pictures of politicians (Figure 28). The Unidad Nacional Española page removed had by far the biggest reach with about 700,000 followers. Facebook took action against these networks after the activist group Avaaz uncovered them and presented their evidence to the company on 12 April 2019. Avaaz's campaign director Christoph Schott said Facebook had done a great job in acting swiftly to take down the pages, however, what Avaaz uncovered was "likely just the tip of the disinformation iceberg". In another research inquiry, Avaaz discovered that around 9.6 million registered Spanish WhatsApp users had received hateful memes and disinformation on the platform in relation to the upcoming election. This figure is apparently higher than the disinformation reach on any other platform as about 89% of Spaniards use WhatsApp. Most of the content disseminated seems to have originated from right-wing extremist groups. Meanwhile WhatsApp took action against the left-wing party Podemos, which had been using WhatsApp as a channel to reach tens of thousands of followers to deliver campaign messages. WhatsApp took the channel offline in the week before the election stating that Podemos was breaking the terms of usage by sending automated mass messages. Podemos said that they were indeed doing exactly that but felt singled out as they are not the only party employing the social media platform in that way.

Podemos is not entirely wrong: generally, Spanish politicians, parties and the government have all learned to utilize the Internet and especially social media platforms. Among the most widely used are Facebook, Instagram, WhatsApp and Twitter. Most parties use these platforms to advertise for themselves and their programme, or, if the party is in power, for the government and their plans. The right-wing extremist party Vox is by far the most successful on Instagram where they regularly share memes and mock other parties (Figure 29), while left-wing extremist party Podemos fares the best on Facebook with over a million likes. Interestingly, Twitter is less influential in Spain and not really actively used by any party. Nevertheless, Podemos has the largest following on the platform.

Figure 25: Post showing a picture of a man who was allegedly hurt by the police in Catalonia. In reality the picture was taken in 2012 in relation to protests by miners in Madrid.

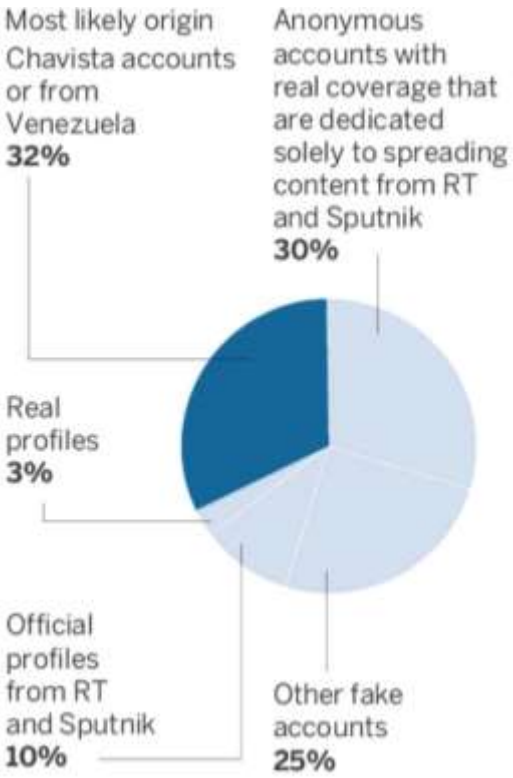
Source: The Washington Post (https://www.washingtonpost.com/news/worldviews/wp/2017/10/19/how-fake-news-helped-shape-the-catalonia-independence-vote/)

Figure 26: Example of disinformation by Russia, uncovered by the EU disinformation task force



Source: https://euvsdisinfo.eu/report/the-logical-answer-from-europe-to-the-catalonia-referendum-would-have-been-recognize-the-independence-of-catalonia-and-bomb-madrid/

Figure 27: Types of accounts used to spread misinformation during the Catalan referendum as analysed by Lesaca



Source: https://elpais.com/elpais/2017/11/11/inenglish/1510395422_468026.html)

Figure 28: Doctored images of Podemos Leader doing the Hitler Salute

Doctored images of Podemos leader Pablo Iglesias doing the Hitler Salute

Caption:"the insults and the lack of support that Podemos is gaining lately and after the celebration of the European elections, are the clear evidence of the LOW level of their crazy proposals, absurd arguments and wrongful conduct"

Source: https://avaazimages.avaaz.org/SpainSummary.pdf

Figure 29: Example Instagram post by right-wing extremist party Vox

Source: *https://www.euronews.com/2019/04/26/weekend-long-read-social-media-use-in-spain-s-election-campaign-the-good-the-bad-and-the-u*

SUDAN

Sudan's media landscape is tightly controlled by the state, and has had a long history of protests – as a result, social media has become a key battleground. Internet penetration is low: of its 40 million population, 13 million use the Internet but more than 28 million own a mobile phone (Reuters, 2019). There have been multiple protests in Sudan since the regionwide Arab Spring in 2011. Most significantly, political uprisings began in December 2018 across several cities calling for President Omar al-Bashir, who has been in power since 1989, to step down. On 11 April 2019, the military removed al-Bashir from power, leading to a transitional government. The political tensions in Sudan have also led to one of the longest Internet shutdowns in history (lasting from June 3rd to approximately July 9th 2019 (netblocks.org, 2019). Leaked documents also show how al-Bashir's government received proposals from Russian firms to quell the protests that erupted (Lister, Shukla and Elbagir, 2019).

The mobilization power of social media led to the creation of the Cyber Jihad Unit, under the National Intelligence and Security Services (NISS), during the wake of the Arab Spring in 2011. The BBC reports that Sudan's ruling National Congress Party (NCP) warned that its "cyber-jihadists" will "crush" Internet-based dissent (BBC, 2011). This unit has received training in Malaysia and India, and is alleged to have 200 full-time employees (Lamoureaux and Sureau, 2019). While the media have popularized the phrase 'cyber jihadists', Kadoda

and Hale (2015) state that the technical specialists are locally referred to as *garad* (grasshoppers). Freedom House notes that, as of 2012, this group is referred to by activists as 'Electronic Chickens'.

The Cyber Jihad unit appears to focus on combatting domestic dissent. It disseminates misinformation to thwart protesters – for example, claiming that protests are a deliberate ploy to destabilize Sudan – and spreads propaganda about the government's handling of the economy (Freedom House, 2018). The unit proactively monitors content on "blogs, Facebook, Twitter, and news and public forums like Sudanese online, Sudan for All, Hurryat and Elrakoba" (APC, 2014). During the 2018 protests, the unit reported pages en masse in an attempt to shut them down. It is claimed they would barrage Facebook pages with pornography and then report them to Facebook for violating their terms and conditions (McClatchy, 2018). In the lead-up to the January 2011 protests, supporters of the NCP posted on activist Facebook pages to warn them against joining protests (BBC, 2011). In one incident on 30 January 2011, the unit allegedly staged a fake demonstration called 'Protesting Youth for Change,' during which 100 would-be protesters were arrested before the demonstration had even started (Lamoureaux and Sureau, 2019). Targeting dissidents and activists is often achieved through social media, by infiltrating online discussions to gather intelligence on online dissidents. Over 15 female activists had their identities exposed on the fake Facebook group 'Sudanese Women against the Hijab' alongside fabricated quotes, which led to violent threats from religious fundamentalists (Freedom House, 2018).

The NISS is said to control digital media use through "blocking, controlling, jamming and slowing down certain websites, and hacking private accounts" (Lamoureaux and Sureau, 2019). NetBlocks, a digital rights organization, said it had found evidence of "an extensive internet censorship regime" (Reuters, 2019). Citizen Lab ascertained through leaked documents that Hacking Team, an Italian offensive cyber weapons company, had at one point sold sophisticated computer spyware to Sudan's government (Citizen Lab, 2014).

Mis- and dis- information also pose problems in Sudan. The government has used misinformation as an excuse to clamp down on opposition. In March 2014, the government declared certain websites would be banned for spreading false information about government activities. In April 2017, during Sudan's fuel crisis, the finance minister stated that he held WhatsApp responsible for spreading false information and creating the fuel crisis (Freedom House, 2018). Given this context, the National Assembly passed the Law on Combatting Cybercrimes in June 2018, making the spread of news online illegal (Freedom House, 2018).

SWEDEN

Sweden is a parliamentary democracy with fair and free elections and a multiparty system. The country has one of the most robust freedom of information statutes in the world which is well respected by government authorities. The media are free and independent and mostly privately owned and state subsidized, regardless of political affiliation. Additionally, public broadcasters regularly air programmes in several minority languages. While threats to journalists have reportedly increased in recent years, such incidents do not seem to

impact the news media's work or lead to self-censorship. Recently, the government decided on new antiterrorism measures which focus on tighter security in public places, greater information sharing between governmental agencies, and better surveillance of individuals deemed to pose a security threat. Most parties are maintaining a social media presence on major platforms (Twitter, Facebook). An interesting and somewhat comical incident occurred in April 2019 when hackers hijacked the Twitter account of the ruling Social Democrats party and announced that Bitcoin would be the new currency of Sweden. They even changed the Twitter name of the party to Bitcoin Democrats (Figure 30). It is not clear what motive lay behind the attack, the Social Democrats quickly regained control of their profile.

In September 2018, Sweden held general elections and, as with most European countries, fake news, trolls and bots flooded the Internet during campaign season. For example, research by the Swedish Defence Research Agency found that the number of Twitter bots increased significantly in the weeks leading up to the elections. Moreover, these bots were 40% more likely to support the anti-immigrant Sweden Democrats than human Twitter users. Troll attacks have also become a common sight online, most of which stem from right-wing extremists which specifically hold immigrants accountable for crime in Sweden. Even foreign alt-right pages have picked up on this narrative such as the American websites Breitbart and Infowars, which describe a neighbourhood in Malmö mostly inhabited by immigrants as a place "ethnic Swedes dare not tread" trying to paint the area as a crime-ridden ghetto. In reality, while it is true that increasing crime rates are a problem in Sweden, these are a result of much more complex reasons than ethnicity, and immigrant neighbourhoods are usually safe and clean areas. Nevertheless, Twitter bots have pushed this anti-immigrant viewpoint since July 2018. The far-right party Sweden Democrats is profiting the most from this narrative and were able to increase their vote share by 4.6% to 17.5% in the elections (Social Democrats received the most votes with 28%).

Alongside the Sweden Democrats, several hyperpartisan news pages, namely Samhällsnytt, Nyheter Idag, and Fria Tider, are profiting from the Twitter bots. These pages are alternative or partisan news' sources, known to share false information. They are read by up to 10% of the Swedish population on a weekly basis. Importantly, it is not clear where the Twitter bots pushing these alt-right topics originate from. Research carried out by the Oxford Internet Institute could not come to any definitive conclusions. Some government officials have considered the usual foreign suspect, Russia, however, many academics believe that the bots originate from within Sweden due to their high fluency in Swedish. Either way, bots are contributing to the wide spread of disinformation and media trolls in the country; however, these two elements have been part of Sweden's media landscape for quite some time. In 2012, the first tabloids and newspapers (such as *Expressen*) started making deliberate efforts to set themselves apart from fake news reports. Meanwhile, trolls have started gaining momentum as citizens are paying increasing attention to societal grievances and are thus more easily seduced by the typical right-wing 'immigration problem' narrative. These stories are also disproportionally pushed by bots. At the same time, politicians and politically active citizens with a more tolerant orientation are targeted by trolls to shut down any form of open political debate.

In light of the elections in September 2018, the Swedish government launched various measures to counter fake news and trolling in relation to the election: (1) a Facebook hotline to report forged pages and profiles to, (2) a "defence authority" to counter disinformation and boost resilience against fake news in the Swedish population by preserving an open society and free exchange of knowledge and information, (3) introducing "source criticism" courses to the curriculum in middle and high school classes, and (4) distributing leaflets with guiding information on how to spot disinformation. Importantly, the afore-mentioned defence authority has only been talked about by politicians, and the Swedish security agencies and the Civil Contingency Agency were mainly tasked with safeguarding the election. Moreover, it is not clear how successful these measures were – they still form part of a uniquely coordinated effort by a European government to combat fake news.

At the same time, Sweden's citizens have started their own initiatives. Between 2014 and 2015, Swedish TV aired a series called *Trollhunters*, a reality show following journalist Robert Aschberg as he tracks down anonymous Internet trolls and confronts them about their actions. The series has been picked up for a new season. Private citizens also started debunking false news on Reddit and Quora and founded a Facebook group in 2016 with around 75,000 members, mainly from Sweden, who under the hashtag #Jhärhär (#Iamhere) defend people attacked online by trolls and try to counter the spread of disinformation in an attempt to rebalance online discussions. While some critics have called #Jgärhär censorship, initiators say they have no agenda and primarily want to spread positivity and love. Since 2016, the concept has spread to other countries in Europe (e.g. Germany and Slovakia).

Figure 30: Tweets sent from the Social Democrats' Twitter account while they were hacked and renamed Bitcoin Democrats

Source: *CNN (https://www.ccn.com/hackers-hijack-sweden-political-twitter-bitcoin)*

SYRIA

Computational propaganda in Syria must be viewed against the backdrop of existing domestic repression, tight Internet controls, and the ongoing civil war. 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 has grown (Economist, 2000). The rule of Hafez's son, Bashar al-Assad was challenged in 2011 by the Arab Spring protests and, following violent repression by the government, these protests have transformed into a complex and brutal war involving both regional and international actors (BBC, 2018). The eight-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 as the Assad regime has long attempted to assert total control over political communication (Freedom on the Net, 2017).

An early report on Syria's cyber troops claimed that in 2011 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 this campaign to a Bahraini company, EGHNA, which successfully flooded the #Syria hashtag in 2011. EGHNA's typical fee for such projects is about US$4,000 (EGHNA, 2017). Katina Michael wrote in *The Conversation* that, in response to Arab Spring activists using hashtags such as #Syria, #Daraa, and #Mar15, government intelligence officers began to threaten online protesters. Syrian blogger Anas Qtiesh wrote that accounts were "believed to be manned by Syrian mokhbarat (intelligence)" with "an endless arsenal of bite and insults" (Michael, 2017). A further tactic was drowning out protesters' voices on Twitter with irrelevant information, for example, photography using #Syria from accounts such as @LovelySyria and @SyriaBeauty.

More than two hundred 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 because activists use mobile phone videos to document the human toll of the conflict, leading to projects such as *The New York Times*' effort to sort, verify, and contextualize videos coming from the region (Stack, 2018). The Facebook pages of dozens of opposition and media groups have been suspended, which activists believe is the direct result of mass-reporting of these pages for violating community guidelines by pro-Assad supporters (Freedom House, 2018).

One 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 registered its domain in 2011 on servers maintained by the Assad-linked Syrian Computer Society, further suggesting that there are tacit links or government support (Freedom House, 2018). The SEA combines cyberattacks and

propaganda by using, for example, phishing to take over the social media accounts of Western news outlets (Harding and Arthur, 2013). In 2013, the SEA hacked the official Associated Press Twitter account and tweeted that Barack Obama had been injured in an explosion, which lead to a momentary panic knocking the stock market value by US$136 billion (Fisher, 2013). Pro-Assad activists reportedly earn around US$500–US$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. Two of the SEA's chief operators, Ahmed al Agha and Firas Dardar, have even made it onto the US FBI's most- wanted list (Forbes, 2018). In May 2018, Open Canada reported that the SEA was re-launching with a new mission as "domestic cyber police", consistent with the Assad government's objective to reimpose sovereignty over the Syrian population (Open Canada, 2018). Further, Forbes reported that the group is putting resources into developing spyware, having developed malware dubbed 'SilverHawk' to target security- and privacy-focused communication apps such as WhatsApp and Telegram (Forbes, 2018).

Each side in the civil war 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). Regional powers (such as Iran) and global powers (such as the US and Russia) have been accused of spreading computational propaganda in the conflict (Di Giovanni, 2016; Cockburn, 2016). 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). On 25 August, the Russian Ministry of Defence released a statement claiming to have intelligence that Syrian rebel forces were about to gas their own people in Idlib province as part of a 'false flag' operation to frame the Syrian government, a claim which was shared by Russian media outlets and supporters of Bashar al-Assad (Telegraph, 2018) (See this report's profile on Russia).

The Syrian Civil Defence, commonly known as the 'White Helmets', are subject to frequent disinformation campaigns. Conspiracies which are promoted by suspicious accounts identified in the BBC study – such as the idea that the Syrian chemical attacks are a hoax created by the White Helmets – are often widely shared by Russian state-run media outlets such as RT and by Western far-right activists. 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). While it was the White Helmets that documented the chemical attack in Khan Sheikhoun in April 2017, which killed at least 83 people, they are continually discredited by an online network of activists, conspiracy theorists, and Russian government trolls (Guardian, 2017). Graphika discovered an online network of 14,000 Twitter users talking about the White Helmets that looked "very similar" and included many known pro-Kremlin troll accounts (Guardian, 2017). Even Russia's official channels post memes discrediting the organization, and alleging they stage 'hoax'

chemical attacks (Figure 31). Investigative journalist agency Bellingcat observed that, from August to November 2018, the Russian and Syrian governments and state-controlled media outlets repeated narratives about the White Helmets and their involvement with chemical weapons in the rebel areas of north-western Syria (Bellingcat, 2018). While the White Helmets have successfully recorded human rights' abuses by the Syrian government, there is yet to be any verifiable evidence that supports the accusations made against them.

In April 2018, there was a surge in disinformation following the chemical attack in Douma. It was claimed that nearly half the counternarrative accounts created in the week between the Douma chemical attack and the Western strike against Syria were inauthentic actors. Sky News reported that the UK government had documented more than 45,000 posts propagating disinformation narratives in the two weeks following the chemical attack on 7 April 2018 (Sky News, 2018). According to BBC Trending, in the hours after the attack, 'Syria' was the top trending term on Twitter, but the messages by pro-Assad activists were overwhelmed by reports from news outlets. The hashtag #SyriaHoax was used around 17,000 times a week in April 2017 but failed to make Twitter's list of top trends (BBC, 2018). Following the US-led missile strike on Syrian targets, the Pentagon claimed a 2,000% increase in Russian troll activity on social media, part of a campaign to present alternative narratives to sow doubt about the evidence that Assad was responsible for the chemical attack (USA Today, 2018).

To combat this disinformation, Ahmad Primo, an activist and journalist, founded the online platform 'Verify-Sy' to monitor and fact-check stories about the Syrian conflict. Their website notes that Syrian 'alternative media' institutions broadcast a huge volume of news and photos, but do not follow journalistic standards in verifying the news, leading to the circulation of fake and inaccurate news content (www.verify-sy.com).

Aware that political support in the United States for a military presence in Syria does not have a 'blank check', the government of Bashar al-Assad has developed and encouraged messaging which undermines the Operation Inherent Resolve (OIR) anti-ISIS campaign. At the same time, pro-SARG (Syrian Arab Republic Government) social media accounts have boosted the role of the Syrian Arab Army (SAA) countering ISIS. Overall pro-SARG social media accounts are promulgating a narrative which depicts the government of Bashar al-Assad fighting a foreign-backed counterinsurgency which seeks to conquer territory which is historically of special importance to the Muslim world. This is being conducted by accounts on Twitter such as "Ivan Sidorenko", (@IvanSidorenko1), "Peto Lucem", (@PetoLucem), "PartisanGirl", (@Partisangirl), and "The'Nimr'Tiger", (@TheNimrTiger or @Souria4Syrians) (O'Leary & Heras, pp.76–78, 2019).

These narrative contestations are components of a broader, multi-domain proxy war in Syria in an era of peer competition. The social media reverberations into and out of Syria geographically may have implications for the future deployments of cyber troops as proxy conflicts accelerate. Social media troops are one element of this approach in the Syrian conflict which includes Electronic Warfare (EW), Information Warfare (IW), and a contested multinational airspace (McCleary, 2018). US Lt. Gen. Paul Funk stated in late 2018 that adversaries such as Russia "want to take us on in the edges, in the information space,

or in EW" (McCleary, 2018). Some researchers suggest that this 'edge' might also extend outwards into exploiting vulnerabilities to create discord in domestic civilian social media environments of militaries, no matter the contribution size. This may even extend geographically to populations as far away from Syria as Australia. In a report to the Australian parliament in late 2018, Tom Sear suggested Russian active cyber measures coincided with a possible correlation with Russian state-sponsored IRA sock-puppet activity influencing opinion on issues like 'syria', 'aleppo', 'merkel', 'isis', even 'sports,' in the Australian Twittersphere when assets were in conflict in airspace over Syria in 2017 (Sear, 2018).

Figure 31: The Russian Embassy supporting White Helmet conspiracy theories

Source: Twitter

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 special, 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 UK (Freedom House, 2017).

Likewise, 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" have been reported in several Southeast Asia counties, 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 fake new followers where 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 centre, 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 $4,403 USD) 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

Turkey continues to experience increasing efforts in computational propaganda. It has been an important year for Turkey, following presidential and parliamentary elections on 24 June 2018. In April 2018, President Recep Tayyip Erdoğan called for early elections, which were brought forward from their November 2019 schedule as Turkey needed to "overcome uncertainty". Following electoral victory, President Erdoğan and the ruling AKP (Justice and Development Party) remained in office with sweeping executive powers that had been narrowly approved in a constitutional referendum in April 2017.

The state of emergency ended in July 2018, which had been in place since an attempted coup in July 2016. The state of emergency resulted in weakened parliamentary and constitutional checks on executive decrees issued by President Erdoğan and his cabinet. One such decree, passed in August 2017, was Decree No.671 which amended the Law on Digital Communications to authorize the government to take "any necessary measure" on the grounds of "national security, public order, prevention of crime" and obliging telecommunications' providers to enforce government orders within 2 hours of receipt

(Freedom House, 2018). While this has been applied as a means of repression, it has also been used positively, such as to enable security agencies to crack down on Islamic State communications.

Social media monitoring increased after the July 2016 coup attempt. 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 citizens' reports of "terror propaganda". According to the Interior Ministry, security forces undertook investigations of almost 50,000 social media accounts for sharing 'terrorist' content online, resulting in "20,000 legal actions taken" (Freedom House, 2018).

Turkish computational propaganda efforts have increased during periods of political uncertainty. Following the 'Occupy Gezi Park Protests' in 2013, the ruling AKP government launched 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 to counter anti-Erdoğan opinions (Guardian, 2016). The pro-AKP *Star* reported that there would be AKP social media representatives in over 900 districts and 1,000 staff located in Istanbul, 600 in Ankara, and 400 in Izmir. In September 2013, the AKP recruited a new social media team, known as the 'New Turkey Digital Office', responsible for converting AKP sentiments into trending hashtags, and engaging in abusive behaviour against journalists and civil society movements (Guardian, 2016). These agencies are believed to have been largely abandoned. Turkey also launched the Directorate of Communications last year. Working directly under the presidency, Fahrettin Altun, Communications Director of the Turkish Presidency, announced that the Directorate will be "at the heart of national and global narrative, consensus, insight and interpretation" and stating their main goal is "to expand the framework of effective communication between the nation and the state based on technology" (Altun, 2018).

There is evidence of automated, coordinated, and inauthentic social media activity relating to Turkish politics. News sources have reported a centralized botnet influencing trending hashtags in two ways: (1) astroturfing (by first boosting a pro-government hashtag), and (2) poisoning (flooding the hashtag with inflammatory content to shade its original message). This was visible in a recent example by the opposition's hashtag, #DemirtaşıÇokÖzledik ("We missed Demirtaş a lot") campaigning on behalf of the jailed pro-Kurdish Peoples' Democratic Party leader, Selahattin Demirtaş, when bots boosted counter-messages to demoralize his supporters (Sozeri, 2017). Similarly, following the 'Gas for Gold' corruption scandal in late 2013, AKP Trolls adopted 'cyber lynch mob' tactics to silence opposition (Okun, 2017). Leaked emails in October 2016 evidenced the government discussion of "a team of professional graphic designers, coders, and former army officials who received training in psychological warfare" to complete these tasks (Freedom House, 2018).

There is evidence that some of these Twitter followings have been created organically. Twitter accounts with large 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 election monitoring group 'Vote and Beyond' (Sozeri, 2015). Similarly, an account under the name 'irem_cevikk' became 'Vote and Fraud'. This fake account used content amplification strategies, using a follower-boosting Twitter application which automated a follow-back system. Another pro-Erdoğan account had 182,000 followers but only nine tweets, and one-month prior had been posting jokes to gain followers (Sozeri, 2015). Bloomberg reported that researchers had found a collection of nearly 18,000 pro-Erdoğan Twitter accounts that used profile pictures taken from pornography sites or public figures such as American actress Megan Fox to gain followers.

The AKP is not the only political party in Turkey using online propaganda. In the build-up to the 2018 elections, the *Hürriyet* daily newspaper reported that opposition İYİ ('Good') Party started a Google Ads' campaign for several keywords targeting the AKP. A Google search for 'vacant rooms' took users to the İYİ election vow to open Erdoğan's palace to the public, and 'VPN' led to the phrase "for Internet freedoms, wait until we come to power" (Hurriyet, 2018).

Online propaganda and repression also support Turkish military policy. In January 2018, there was a wave of arrests in response to the critiques of 'Operation Olive Branch', a Turkish military operation in Afrin, Syria. Turkey asked social media sites, such as Facebook and Twitter, to take down posts that criticized the operation. In 2017, Turkey ranked as the country with the highest number of content removal requests sent to Twitter, with 4,294 requests filed by Turkish government agencies, according to Twitter's transparency reports (Twitter, 2018). The Turkish Interior Ministry claimed that they detained 648 people between 20 January and 26 February 2018 over social media posts criticizing the military operation (Human Rights Watch, 2018). People were arrested for "posting information on social media from local sources in Afrin that contained alternative rhetoric to that of the government". Journalist Nedim Turfent, who was reporting on counterterrorism in Turkey's Kurdish region, published a video of soldiers standing over villagers who were face down with their hands bound. Messages seeking Turfent's whereabouts began to appear on his Facebook page and Twitter accounts linked to Turkish counterterror units began to taunt locals with "Where is Nedim Turfent?". Within days, Turfent was detained by the military and charged with membership of a terrorist organization (Bloomberg, 2018).

Non-government actors that support government policy are also active in propaganda efforts. A hacking collective called Ayyildiz Tim (the 'Turkish Cyber Army') has increasingly peddled pro-government messages on Twitter through hacking prominent figures' accounts. There are multiple 'Turkish Cyber Army' groups and these are active on social media (Figure 32). Since summer 2017, the TCA has focused on Twitter phishing to compromise accounts and, upon gaining access, has made pro-Turkish posts, downloaded message history, and sent new phishing attacks. The group has even managed to direct-message US President Donald Trump on Twitter, after having gained access to the Twitter account of the head of the World Economic Forum, Børge Brende, who is followed by Trump. After an account is hacked, its Twitter bio would typically read: "Your account has been hacked by the Turkish cyber army Ayyildiz Tim. Your DM correspondence and important data have been captured!". Chuka Umunna, a British Member of Parliament, was

hacked in March 2018, with his compromised Twitter account posting references to the Turkish military operation in Syria (Figure 33) (Evening Standard, 2018). While there is no immediate evidence that the group is a Turkish government organization, they are strong supporters of the government. The Turkish government launched an official cyber army to defend against cyber threats. In April 2017, the Ministry of Interior and Information Technologies Directorate jointly announced the establishment of a 500-strong cyber army (Haberturk, 2017).

Online harassment and repression targets Turkish journalists. In 2018, Reporters Without Borders ranked Turkey at 151 out of 180 countries in their Press Freedom Index. Fake news is prolific, with the Reuters Institute's 2017 report finding that 49% of Turkish people had been exposed to 'fake news' within the previous week, and 38% said they did not trust the news (Reuters Institute, 2017). Individual targeting of journalists often consists of 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. Trolling was turned into real lynching in 2016, when the *Hürriyet* newspaper building was attacked by protesters. Female journalists are most often targeted by hundreds of trolls with sexual-related insults, such as the case of Nevsin Mengu, a popular CNN-Turk anchorwoman who was forced out of her job following her coverage of the 2016 coup attempt (Institute for the Future, 2018).

Fact-checking organizations have been founded in response to the manipulation of social media, tasked with attempting to combat disinformation. Mehmet Atakan Foca, the editor-in-chief of Teyit.org (Turkish for 'confirmation') said that the organization receives 25 to 30 reports of suspicious messages, images, and videos every day (Edroos, 2018). Fact-checking portals themselves are becoming tools of government propaganda and misinformation. A BBC report (2018) found that even fact-checking itself is being used as a tool to sow mistrust and division. One website claims to be an independent verifier of news, but is in fact run by a prominent columnist for Sabah, the main pro-government daily. 'Factcheckingturkey.com' is an English-language fact-checking site that aims to check foreign media coverage of Turkish news (Figure 34), which is said to have been founded because Turkey was being "represented as yet another dictatorship" in the foreign media. Similarly, another fact-checking site is run by a PR firm, Bosphorus Global, which aims to verify news but has close ties to the government (Politico, 2019). This tendency extends outside Turkey too – for example 'factcheckarmenia.com' denies the Armenian Genocide of 1915, and is active in the United States but tied to Turkish government-affiliated organizations.

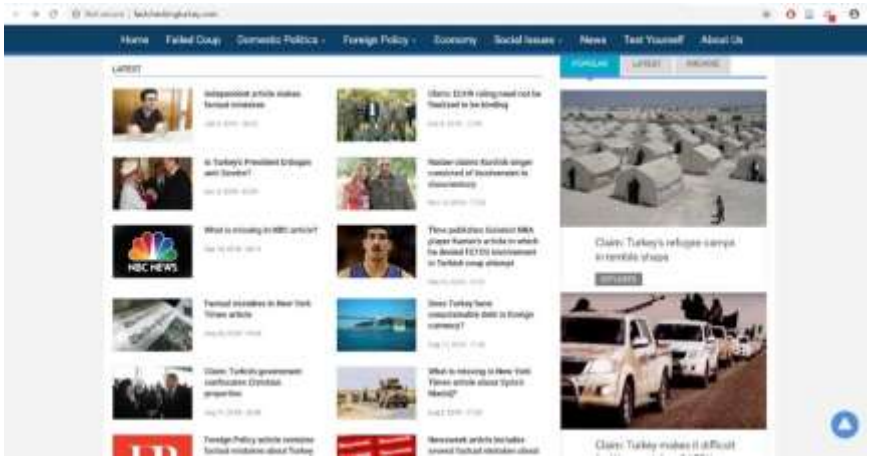
Figure 32: Ayyildiz Tim's Twitter bio reads: "When one of us dies, we resurrect in thousands. Cyber Army of the Turks. It's not over until we say it's over."

Source: Twitter, 2019

Figure 33: Chuka Umunna's hacked Twitter account

Source: Evening Standard, 2018

Figure 34: A fake fact-checking portal



Source: *http://factcheckingturkey.com/*

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 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 organisations 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. organised 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 percent of respondents believe in pro-government manipulation in online debate (Internews, 2017). Furthermore, both the ministry and the i-army initiative were widely criticised 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

The United Arab Emirates' (UAE) efforts at computational propaganda form part of a coordinated military and public diplomacy effort. Efforts are complemented by a range of harsh domestic social media and cybercrime laws, which criminalizes offending the state, its rulers or symbols, insulting religion and 'sympathy for Qatar'. For example, human rights activist Ahmed Mansour was sentenced to 10 years in prison in May 2018 for "spreading sectarianism and hatred on social media" (Freedom House, 2018).

The UAE's propaganda efforts cannot be examined in isolation from other initiatives, like the funding of think tanks and conventional media in a campaign to disseminate narratives favourable 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 the 2018 report's interview with 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 political matters in the country, 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 research funding activities of the UAE, Krieg compared Emirati efforts to those adopted by Russia, albeit unsuspected, being predominantly under the radar.

Emirati computational propaganda reportedly started out as part of a defensive strategy in late 2012, spreading positive messages about the UAE, targeting mainly US and, to a lesser extent, UK audiences. In the context of the 2011 Arab Spring, the UAE deployed a more aggressive strategy including foreign attacks on political Islam (conflating any form of

political Islam with Islamic State-type Salafi-jihadism); against Qatar, Turkey and Iran; but also against opposition forces in Syria, Iranian proxies in Iraq (such as Islamic Revolutionary Guards Corps section outside Iran), and the Egyptian grassroots movement Tamarod. 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 example, orchestrated campaigns to discredit 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 favour of the Gadaffi regime. This strategy was further refined in the context of the 2014 Gulf Crisis, especially since advancing national interests online seemed to have become a convenient tactic in comparison to other 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.

In May 2017, the Qatar News Agency (QNA) was hacked, and remarks – attributed to the Emir of Qatar – were published in support of Iran, Hezbollah and Hamas, and critical of Donald Trump. These were further relayed by Emirati and Saudi news channels and spread on social media. A US intelligence investigation revealed that the UAE orchestrated the hacking of the QNA and its social media sites in order to post incendiary and false quotes attributed to the Qatari Emir to spark a divide between Qatar and its neighbours (Washington Post, 2017). It is claimed that the hack was the deed of Russian hackers hired by the UAE, but the UAE has denied attribution. Jassim Al Thani, Qatar's Washington-based media attaché, stated the UAE "weaponized fake news to justify the illegal blockade of Qatar" alongside using "cyberespionage, fake news and propaganda" (Buzzfeed, 2018).

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. According to Andreas Krieg, Abu Dhabi has created a powerful web of policymakers, think tanks and experts in the United States, aligned to neo-con and AIPAC (the American Israel Public Affairs Committee) positions, influencing Washington discourse a positions on Middle Eastern affairs, as well as shaping the Trump administration's initial approach to the region. This emerged in the Mueller investigations over UAE's attempts at buying influence during the 2016 presidential campaign, however, appears to have been played out in person rather than online. This took the form of visits and lobbying, as businessman and adviser to the Abu Dhabi Crown Prince George Nader often visited the White House and met with Senior Advisor Jared Kushner and former Chief Strategist Steve Bannon.

Since 2014, propaganda efforts have expanded to the West, and campaigns outside the Arab world have been outsourced to public relations and consulting firms in the US, the UK, Germany, Switzerland, and many others, as the Emirates lacked the propaganda capability or capacity. PR and lobbying firms worked "to sway American public opinion through online and social media campaigns" (PRI, 2018). The Huffington Post reported in 2015 that the UAE had spent more than US$12 million on lobbying and PR from 2014 to 2015. Many suspect that these firms are used to counter negative images of the country's human rights abuses online, as a large number of anonymously operated Twitter accounts appear

dedicated to harassing political dissidents (Freedom House, 2018). This social media manipulation was particularly evident during the Gulf Crisis in 2017, in which Gulf Cooperation Council (GCC) states, including the UAE, cut diplomatic ties with Qatar, following the false news reports attributed to the UAE hack of the QNA. In 2017, Cambridge Analytica executives created Emerdata, under parent company SCL Social Ltd, which was reportedly awarded a US$330,000 contract from the National Media Council of the UAE for social media outreach. The company is recorded as spending US$60,000 on ads on Facebook, YouTube and Twitter to promote the #BoycottQatar hashtag, and links to articles critical of Qatar alongside disinformation (Medium, 2018). The Harbour Group, which has represented the UAE for over 15 years, was allegedly paid more than US$2.5 million by the UAE for work between October 2016 and March 2017.

Social media has also been used for surveillance efforts. From as early as 2015, the Telecommunications Regulatory Authority has been monitoring social media networks, including an automated "alert system that will detect when certain keywords are being used" (Arabian Business, 2015). In February 2016, an official from Dubai Police said authorities monitor users on 42 different social media platforms (Freedom House, 2018). In December 2018, Dubai Police reported that they had blocked 5,000 fake social media accounts in the UAE through a "smart system", an "automated system that monitors this type of account" (MSN, 2018). Surveillance is aided by private cybersecurity firm DarkMatter, which acknowledged that 80% of its customers are UAE government agencies. One former DarkMatter operative, and former NSA employee, stated that under order from the UAE government they would monitor social media and target people who security forces felt had insulted the government (Reuters, 2019).

The UAE currently controls a wide range of traditional news outlets, including their respective social media presences: al-Arabiya (the network is Saudi but operates from Emirati capital Abu Dhabi), which has frequently denounced Iranian and Qatari attempts at computational propaganda, and Sky News Arabic. In its endeavour to create a new narrative as a tolerant Middle Eastern partner that shares US security concerns and spreads anti-Qatar messages, the UAE engaged with the US in the creation of the Sawab Center in 2015. This is dedicated to countering the so-called 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, the Center's aim is to "amplify moderate and tolerant voices from across the region". Since then, it has 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 to uncover violent extremists' lies. 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 aimed at supporting the UAE in its efforts to combat radicalization and become the champion of counterterrorism in the Arab world.

UNITED KINGDOM

The UK has an Internet penetration rate of 94.8% with generally free access. There are no known incidents of governmental control over content or any amount of Internet data transmitted across the country, and its infrastructure is well developed. The UK High Court can order the blocking of websites, but only in cases of copyright infringement. In recent years two laws have caused public controversy: the Investigative Power Act, which passed into legislation in November 2016; and the Digital Economic Act, which passed in April 2017. The former deals with, among other things, interception, equipment interference and data retention, while the latter was supposed to redefine the terms under which "extreme" pornographic material can be blocked online, as well as introduce age verification methods aimed at protecting minors from online pornography. In January 2018, the Data Retention Provision of the Investigative Power Act (DRIPA) was ruled inconsistent with European law by the UK Court of Appeals, and the High Court ruled parts of it unconstitutional in April 2018, requiring politicians to amend the DRIPA by November 2018. According to a government-issued statement these amendments were made and mainly consist of introducing stricter judicial oversight for law enforcement to access certain powers granted under the Act. Also, the age verification requirements under the Digital Economy Act are still cause for concern: the act was supposed to become active in April 2019 but was postponed again. Critics, such as Jim Killock, Executive Director of the UK's Open Rights Group said: "the policy is full of holes". He asserts that the age restriction is doomed to fail and liable to do more harm than good.

Meanwhile, a wealth of academic literature and journalistic material has dealt with the presence and extent of computational propaganda in recent elections in the UK – including during the 2017 general election, the 2016 EU Referendum (or Brexit), and the 2014 Scottish Independence Referendum. These findings highlighting the prevalence of computational propaganda led to the government accusing social media platforms of not doing enough against hate speech so, in October 2017, the government announced the establishment of a national hub which will monitor hate speech and support its removal. The initial budget was reported at £200,000 in March 2018.

Automated content has been reported during the 2017 election. Facebook announced in May 2017 that it had removed 10,000 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 reasons for voting 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.
Leaks connected to 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 questioned following revelations around the Leave campaign's outmanoeuvering of spending limits (by donating £625,000, US$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.

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 UK politics came from professional news organizations; junk news accounted for 11.4% of news content shared on Twitter. Additionally, 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 UK. 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.

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 favour of Independence. Automated activity is suspected to have been orchestrated by the Kremlin. In addition, post-referendum where the Remain in the UK 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.

In 2015 the British Army set up the 77th Brigade in an effort to "challenge the difficulties of modern warfare using non-lethal engagement and legitimate non-military levers as a means to adapt behaviours of opposing forces and adversaries", as it says on the Army's website. The new division is assembled from several older parts of the British Army, namely the Media Operations Group, a Military Stabilisation Support Group and a Psychological Operations Group. It seems the Brigade uses both Twitter and Facebook in their information warfare. In early 2019, reports surfaced saying that their Twitter profile (@77th_Brigade) had been hacked by an individual who started trolling the Army's Twitter account for its easy hackability (Figures 35 and 36). Eventually the army was able to regain control of the account, which the hacker had renamed @79th_Brigade and locked it so non-followers could not read its tweets. How valid this story is remains questionable as no major newspaper reported on it, while one minor paper appears to have published the exact some story as *Sputnik News* which is known to be Russian owned.

Figure 35: Newly renamed 79th account trolling Army (allegedly)

Source: *https://sputniknews.com/news/201902221072660490-77th-brigade-twitter-hacked/*

Figure 36: Newly renamed 79th account trolling the British Army (allegedly)

Source: *https://sputniknews.com/news/201902221072660490-77th-brigade-twitter-hacked/*

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 $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 US 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, amongst 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 $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 US-based "extremist" Somalimidnimo.com website, undertaken via US 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 US tradition of public diplomacy through print media and broadcasting, as well as counter-propaganda activities. The US 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 $580 million a year, but this was progressively reduced to $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, US diplomats have begun actively trolling Isis, arguing with pro-Isis accounts and producing videos portraying Isis-conquered territory as a hellscape. US 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 $200 million in funding since. The software it developed, contracted for $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 US-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 US elections, in both camps, relying on interactive advertisements, live-streamed video, memes, and personalized messaging. According 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 astro-turfing 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 unfavourable video or audio of Trump), as well as Shareblue, which with a budget of $2 million was focused on exposing alleged news coverage against Hillary Clinton and extensively engaged in astro-turfing 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 astro-turfing by means of viral videos and memes, which behave like political ads 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 “grass-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.

Astro-turfing 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 US 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.

There have been a number of instances of foreign governments involved in influence operations targeting American citizens. Most of the revelations over the past year have focused on how Russia's Internet Research Agency used social media to sow divisions and discord between American voters and interest groups, targeting groups such as Black Lives Matters, gun right activists, and pro-life groups. However, other countries—such as Iran and Venezuela—have launched campaigns on social media targeting American citizens.

The growing awareness around the efforts of foreign nations, and especially Russia, to influence US 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 $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 US has also supported counter-propaganda efforts abroad, financing during the Obama administration up to $1.3 billion for Europe alone to strengthen resilience against Russian meddling.

Not all social media manipulation is conducted by state and political party actors. In the United States we saw one of the first examples of a so-called 'shallow fake' video go viral with Democratic speaker Nancy Pelosi's voice slowed down to make her appear intoxicated. While this video was not created as part of a coordinated and weaponized disinformation campaign, the way it was picked up and shared by politicians and other media outlets highlights the fact that social media manipulation takes place in a broader ecosystem where many different actors interact with one another.

VENEZUELA

Venezuela is one of the least free countries in the world and gas been suffering from a long economic crisis and a violent authoritarian regime (MrBarbacoa2011, n.d.). Plunged in recession, communication has been weaponized by the government and opposition in a struggle to maintain – or break – control over unsatisfied population.

The country scores 66 in the Freedom on the Net ranking, which is a sign of the permanent threats and ongoing attacks on journalists, media outlets and communication infrastructure. Online propaganda had already been registered in 2010, during the Hugo

Chávez regime (2002-2013), when the then-president announced his usage of Twitter (MrBarbacoa2011, n.d.). The Chávez government produced widescale propaganda, including disseminating YouTube campaigns and summoning people to follow him on Twitter during presidential speeches.

In 2018, an important leak of governmental documents showed how the department of interior was creating a cyber-militia, with structured teams and incentive systems for propaganda dissemination (*A Global Guide to State-Sponsored Trolling*, n.d.). This included instructions to build social accounts, guidelines on the creation of strategy groups, incentives and even the creation of specialized tasks such as content creation, distraction or attack.

Disinformation teams were organized in a military structure, where 23 individuals, organized in "companies", could operate as many as 1150 accounts. People participating in these operations were rewarded with coupons for food and goods, which are particularly valuable in the current state of scarcity (*A Global Guide to State-Sponsored Trolling*, n.d.).

Both the current government and the opposition parties have both been weaponizing internet communication to try to rally support in their power dispute. Since the 2018 elections, Venezuela is in a tug-of-war between Nicolas Maduro, president in office since Chávez passed away in 2013, and Juan Guaidó, who claims the elections were rigged and self-appointed the new president of Venezuela.

In fact, it is not uncommon for political figures to use their communication powers to promote attacks against their opponents through propaganda. For example, in 2016, former vice-president Diosdado Cabello used his TV channels and Twitter accounts to promote a hashtag attacking opposition politician Luis Florido ("IFTF: Government Sponsored Trolling," n.d.).

Disinformation and social media attacks have been present in the countries for years. In 2017, a community of 32 influencers, including DolarToday, a Miami-based website that is said to offer a reliable exchange rate between the USD and the Venezuela bolivar, and who accumulated, at the time, 2.9 million followers on Twitter (Gallagher, 2017). It remains unclear what sort of organization was behind this group of accounts, but they orchestrated tens of thousands of retweets and generated millions of views in favor of protests.

Information warfare intensified during the campaign period. Since the election results are still contested, the opposition is using social media to spread false information in an attempt to rally citizens to protest against the government, break morale of government supporters and counteract the intense governmental propaganda machine.

Cybertroop activity is present on most mainstream social media, including YouTube, Facebook, Instagram, Telegram, Google+ (now extinct), WhatsApp and especially Twitter. An automated Twitter account linked to the government incentivized the daily retweeting of governmental hashtags. It rewarded the top supporters periodically with cash prizes. There are also registers of the account orienting individuals on how they could participate

in these retweeting activities and win prizes (@DFRLab, 2019a). There has also been plenty of evidence of state-sponsored tolling, where people were hired to attack social media accounts of journalists and opposition supporters(@DFRLab, 2019b).

Tensions have been escalating since the beginning of the year, when Maduro's runner up in the presidential elections challenged the results of the elections. The country is since deeply immersed in a presidential crisis. These tensions led to an uprising on April 30th 2019, where Juan Guaidó claimed to have control over Venezuela's main airbase and support from military generals to take over the Venezuelan government (Smith & Torchia, 2019). During the uprising, false information circulated that Maduro had resigned, while protestors took the streets. Since January, rumors of Maduro resigning have been strategically spread during moments of tension ("Amid Chaos, Venezuelans Struggle To Find The Truth, Online," n.d.). During the period of crisis, the United States sent over humanitarian aid trucks and John Bolton claimed that these rescues were being set on fire by the Maduro government. It was later revealed that the only truck that caught on fire was provoked by the protestors themselves, who were throwing Molotov cocktails at the police and accidentally set one of the trucks of fire ("Footage Contradicts U.S. Claim That Nicolás Maduro Burned Aid Convoy - The New York Times," n.d.).

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 rumours and blocked class 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. The troll army in Vietnam has been known to target and harass Vietnamese activists and civil society organizations.

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 prioritise governmental issues with fake news that circulates as content or as ads.

In light of cybersecurity and fake news concerns, the government drafted a cybersecurity 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 cybersecurity and taking aim at controlling content. It will be voted on by the National Assembly in June 2018. On June 8, 2018, the US 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, who freedom of expression will be curtailed if the government has access to Vietnamese data on social networking platforms.

ZIMBABWE

As with other African countries, mobile phones are the most common way through which Zimbabweans connect to the world. Still, millions of citizens remain unconnected due to poor networks. Most online communication occurs on social media platforms, mainly WhatsApp, Facebook and Twitter for which users can buy bundles for up to 250MB and daily social media access. Internet access is provided by five services, three of which are government-owned, and two privately owned. In July 2018 Zimbabwe experienced its first election where social media played a significant role, with main candidates looking to fund teams dedicated to running their campaigns online. Most Zimbabweans have turned to social media to stay up to date with their country's politics. At the same time, the spread of misinformation has increased significantly, especially during the military intervention leading Mugabe to resign and during the election campaigning time in 2018. Allegedly the ZANU–PF, the ruling party since 1980 (first under Mugabe who was forced to resign in late 2017, now under Mnangagwa who won the election in July 2018) is paying pro-government commentators to defend the administration and attack oppositional parties and critical voices on social media. Additionally, state-owned news outlets have also engaged in attacking opposition parties to the ZANU–PF by spreading rumours about opposition politicians like Chamisa who was falsely reported to have been consorting with Mugabe.

In general, Zimbabwe's government does not censor online content, but rather controls broader access to the Internet and social media: in July 2016, WhatsApp was reportedly blocked during nationwide anti-governmental protests; in July 2018, the independent advocacy organization Zimbabwe Election 2018 was blocked by the state-owned Internet provider TelOne; in early 2019, the Internet was blacked out and social media shut down for days when citizens took to protesting against rising fuel prices during the current economic crisis. In February 2019, rumours started circulating online that a new currency

would be introduced in response to the economic crisis, to which the government reacted by tweeting from their various Twitter accounts (Figures 37 and 38). At the same time, critical voices and independent media have criticised state media for spreading what they call pro-government propaganda by reporting that the government is on its way to restoring Zimbabwe's economy with headlines such as 'Fuel and wheat storage already fixed' (Figure 39).

The current economic situation in Zimbabwe brought to light the controlling attitude of the government towards its own citizens. In January 2019 fuel prices increased by 150% which lead to widespread discontent and violent protests. As a reaction to the protests Zimbabwe's authorities started firing at demonstrators, killing at least a dozen people, shutting down social media, and, as mentioned, eventually ordered Internet blackouts. Initially, the government denied having anything to do with the blackouts, but soon warrants emerged which showed that the state-owned Internet providers were ordered to stop their service by the National Security Minister. Hashtags such as #ShutdownZimbabwe started trending as citizens aired their frustration about the government's continued control over Internet access. International organizations such as Freedom House urged the government to stop the violence and allow for protest both online and offline.

Interestingly, the warrants issued to black out the Internet are legal and were made possible through the Interception and Communication Act introduced in 2007, which enables the government to intercept telecommunication for the purpose of protecting national security. Following the blackout and social media shutdown, a group of lawyers formed calling on the High Court of Zimbabwe to rule the policy as unconstitutional. In late January 2019, the High Court ruled that only the president – and no minister – could legally block the Internet. Future decisions to stop services will have to be made on a case-by-case basis by the president himself. Importantly, social media are playing a crucial role in Zimbabweans finding fuel. In October 2018, the hashtag #FindFuelZW started trending on Facebook and Twitter, with activists such as Kuda Musasiwa actively supporting the efforts of Zimbabweans coordinating online (Figure 40). Moreover, software developers from Intelli Africa Solutions created a chatbot to help find fuel around the capital city Harare, which was launched in late October 2018 and had over 9,000 users in the first month (Figure 41).

Meanwhile, there is growing concern that Zimbabwe is slowly but surely turning into a police state. In 2016 Mugabe introduced the Computer Crime and Cyber Crime Bill which would penalize the dissemination of communications "with intent to coerce, intimidate, harass, threaten, bully or cause substantial emotional distress" with fines and up to 10 years in prison. Critics have said that the bill would mainly restrict the freedom of expression online. In 2018 Mnangagwa announced that the bill would be merged with the Data Protection Bill and the Electronic Transfer and Electronics Commerce Bill. In January 2019 all three bills passed into legislation as the Cybercrime and Cyber Security bill after it was fast-tracked by Zimbabwe's Information Minister. Allegedly, this was done as a reaction to the ongoing protests against the economic situation in Zimbabwe. Critics such as Charlton Hwende (chairperson of the Parliamentary Portfolio Committee on Information Communication Technology), say the bill "lays the foundations of a police state" as it can be used to legitimize the surveillance of government critics and citizens on social media.

In March 2018, the National Policy for Information and Communications Technology passed into legislation, which aims to centralize control over the country's Internet backbone. Officially, it is supposed to foster growth in the ICT sector and eradicate corruption by eliminating bureaucratic bottlenecks. However, the policy seems part of a wider plan to bring Internet Service Providers under government control: in 2016 the Mugabe administration reportedly paid US$21 million to acquire Telecel, one of the five main providers. Licences for the providers are continuing to increase in price, although the government-owned companies do not have to pay the full fee.

In addition, Zimbabwe is receiving support from foreign countries to increase their surveillance capabilities. In March 2018, the country partnered with the Chinese company CloudWalk Technology, which provided a nationwide, public facial recognition software. While Zimbabwe's authorities say the software will be mainly used by law enforcement and security services, other departments may receive access to the programme in the future. Meanwhile, the Japanese ambassador to Zimbabwe delivered a grant of US$3.6 million to the Zimbabwean Finance Minister for cybersecurity equipment. The grant is reportedly part of the wider Japanese "Grant Aid Project for Cybercrime Equipment Supply". Zimbabwe announced it would use the grant to enhance digital forensics and facial recognition.

On a positive note, Mnangagwa's administration dropped the Ministry of Cyber Security, Threat Detection and Mitigation, which Mugabe established in October 2017 and was largely assumed to be used to further restrict online freedom. The *Freedom on the Net 2018* report on Zimbabwe said there were no cases of prosecution for online activity while the report was being drafted. Nonetheless, critical and independent media outlets are usually based inside the country, as local stations are facing threats and arrests by state officials. Thus, self-censorship remains quite prevalent across the country.

Figure 37: Ministry of Information, Publicity & Broadcasting deny currency rumours on Twitter

Figure 38: Reserve Bank of Zimbabwe denying currency rumours on Twitter

Figure 39: Sunday Times headline on government 'fixing' economic crisis

Source: https://www.sundaymail.co.zw/government-dealing-with-economic-challenges-ed-fuel-and-wheat-shortages-already-fixed

Figure 40: Activist Musasiwa supporting the efforts of the hashtag #FindFuelZW

Source: *New York Magazine (http://nymag.com/developing/2018/11/zimbabwe-whatsapp-chatbots-social-media-fuel-shortage.html)*

Figure 41: Chatbot supporting Harare citizens in finding fuel

A text conversation on WhatsApp between a user and Intelli Africa Solutions' chatbot, displaying gas availability at a station in Eastlea, Zimbabwe. Photo: Christine Ro

Source: *New York Magazine (http://nymag.com/developing/2018/11/zimbabwe-whatsapp-chatbots-social-media-fuel-shortage.html)*