Canadian Security Intelligence Service | Service canadien du renseignement de sécurité

WHO SAID WHAT?

The Security Challenges of Modern Disinformation

ACADEMIC OUTREACH

Canada

Think recycling

This document is printed with environmentally friendly ink

World Watch: Expert Notes series publication No. 2016-12-05

This report is based on the views expressed during, and short papers contributed by speakers at, a workshop organised by the Canadian Security Intelligence Service as part of its academic outreach program. Offered as a means to support ongoing discussion, the report does not constitute an analytical document, nor does it represent any formal position of the organisations involved. The workshop was conducted under the Chatham House rule; therefore no attributions are made and the identity of speakers and participants is not disclosed.

www.csis-scrs.gc.ca

Published February 2018
Printed in Canada

Photo credit: gettyimages.com

WHO SAID WHAT?

THE SECURITY CHALLENGES OF MODERN DISINFORMATION

HIGHLIGHTS FROM THE WORKSHOP

TABLE OF CONTENTS

The workshop and its objectives 1

Executive summary 5

Orchestrated or emergent? Understanding online disinformation as a complex system 13

Russia, the West and the geopolitics of disinformation 23

NATO's eastern flank: A new battleground 31

Foreign influence efforts and the evolution of election tampering 41

Examining Brexit: The rise and fall of a Twitter botnet 51

Applying open-source methods to debunk 'fake news' about Syria 59

China's approach to information and influence 69

From 'likes' to leaders: The impact of social networks in the Philippines 79

Countering disinformation in Ukraine 89

Fake for profit: Non-state actors and the business of disinformation 97

Endnotes 105

Annex A: Workshop agenda 111

Annex B: Academic Outreach at CSIS 115

THE WORKSHOP AND ITS OBJECTIVES

On 20 November 2017, the Academic Outreach (AO) program of the Canadian Security Intelligence Service (CSIS) hosted a workshop to examine the strategic impact of disinformation on national security and the integrity of democratic institutions.

Held under the Chatham House rule, the workshop was designed around the knowledge and experience of a multi-disciplinary group of experts from Canada, the United States and Europe. The presentations and plenary discussions allowed attendees to explore the manipulation of information for political and related purposes, examine several recent cases, and critically discuss related security threats. The papers presented at the event form the basis of this report. The entirety of this report reflects the views of those independent experts, not those of CSIS.

The AO program at CSIS, established in 2008, aims to promote a dialogue between intelligence practitioners and leading specialists from a wide variety of disciplines and cultural backgrounds working in universities, think-tanks, business and other research institutions in Canada and abroad. It may be that some of our interlocutors hold ideas or promote findings that conflict with the views and analysis of the Service, but it is for this specific reason that there is value to engage in this kind of conversation.

EXECUTIVE SUMMARY

The reach and speed of the Internet and social media have escalated the potential impact of disinformation. Increases in data transmission capacity coupled with a shift towards programmatic advertising[1] have resulted in a precipitous decrease in the ability of traditional journalism to mediate the quality of public information. Conventional journalism has been partially displaced by a torrent of data from an infinite number of originators. Within that torrent is a current of lies and distortions that threatens the integrity of public discourse, debate and democracy.

Agents of disinformation: The actors

Disinformation has become a highly effective tool for state actors, profiteers, status seekers, entertainers and true believers. The most skilled national purveyor of falsehoods is Russia. Its historic mastery of 'special measures', magnified by modern technology, follows the basic operational principle of *vilify and amplify*:

- Russia's adhocracy, the shifting elite around President Vladimir Putin, directs an extensive network of Internet trolls and bot networks which generate and spread material across the web. Their activities are intensified by the support of diplomats, state-controlled media outlets such as RT (Russia Today) and Sputnik, as well as de facto alliances with organisations such as WikiLeaks;

- Working together, these agents of the Russian state can create a false story and ensure it reaches the segment population most likely to be influenced by it through Facebook, Twitter and other channels. They also appear to corroborate the story through news agency interviews featuring phoney experts, forged documents, and doctored photos and videos. Anyone who challenges the lies becomes a target for high-volume online vilification; and

- Russia, China and the Philippines use disinformation techniques to control their internal populations. Russia stands out for its highly organised strategy of using disinformation to interfere

with the political systems of other countries, influence the political views of its citizens, and create and exacerbate division and distrust.

Both Moscow and Beijing have developed sophisticated information doctrines as part of their strategy to consolidate control domestically, and to advance foreign-policy objectives. Both coordinate messages across multiple platforms, with consistent lines advanced through regular news outlets and social media in many languages. Disinformation serves immediate and longer-term strategic objectives. There are important differences, however, between the Russian and Chinese approaches:

- Russia attempts to alter the perception of reality, and identifies exploitable divisions in its target audiences. It pushes a nationalist agenda more than an ideological one and targets the Russian population to prevent dissent. The surrounding band of states which were once part of the USSR are attacked with messages which may ultimately support hybrid warfare. Operations against Western populations aim to weaken resistance to Russian state objectives. In supporting Syria, Russia has used disinformation to cover the brutality of its attacks on civilian populations;

- China has created a domestic cyber fortress, and reinforced it with Chinese technology and Chinese high-tech companies. The messages projected domestically and globally are both nationalistic and ideological. Beijing uses its version of soft power to influence the policies of the international community, making effective use of economic power and the presence, in countries of interest, of Chinese populations and businesses; and

- Russia's disinformation machinery is explicitly weaponised as a resource for future wars, weakening a target country's sense of danger and diminishing the will to resist. China wants acceptance of its legitimacy as a great power while rejecting international standards it does not agree with.

The stream of disinformation also flows from other actors:

- In the Philippines, disinformation has been used as a tactic to influence voters in the presidential election, justify the street anti-drug campaign, discredit critics, and de-legitimise mainstream media;
- During the Brexit campaign large numbers of Twitter accounts were active, particularly on the Leave side. Most disappeared immediately after the vote, strongly indicating they were driven by bots. In their content they reflected the hyper-partisan and simplistic style of the British tabloid press.

Independent emergent activists

State disinformation agencies are part of a complex system which includes independent activists with different but overlapping motivations. Many see hidden conspiracies behind headline events such as mass shootings, or even deny that they happened. They believe Western governments are untrustworthy, manipulate world events, and are aided in hiding the truth by the traditional media. Most are anti-globalist, with a nationalist and anti-immigration rhetoric that attracts elements of both the left and right.

Independent actors use social media and specialised web sites to strategically reinforce and spread messages compatible with their own. Their networks are infiltrated and used by state media disinformation organisations to amplify the state's own disinformation strategies against target populations. The extent to which activities within this complex system are orchestrated, and by whom, remains unclear.

Agents of disinformation: The enablers

The information ecosystem enables large-scale disinformation campaigns. False news is spread in many ways, but Facebook and Twitter are especially important tools. Both are used to target specific population segments. Individuals accept the false news as credible or useful, and spread it further. State agencies make extensive use of

bots and phoney accounts to popularise false news stories, and spread them in cascading volumes impossible for human actors to produce or vet individually.

Social media companies are becoming aware of their role in the problem, but not all Silicon Valley leaders are convinced of their responsibility to eliminate false news. Fighting spam is a business necessity, but terminating accounts or checking content constrains profitability. Social media companies have a philosophical commitment to the open sharing of information, and many have a limited understanding of the world of intelligence operations. They are reluctant to ally with intelligence agencies and mainstream news organisations to take up the detailed task of monitoring content.

Russian disinformation: The messages

Russian disinformation is adjusted to circumstances and state objectives, but there are persistent major themes according to which, for example, Western governments are fascist, or world leaders represent a powerful elite disdainful of, and acting against, ordinary people.

To these general themes are added those which support specific campaigns, such as Russian activity to support the Republican Party during the 2016 presidential campaign in the United States.

The reaction

Multiple actors and agencies are working to counter and defend against this threat:

- Governments are increasingly insisting that social media companies take responsibility for the content they facilitate. European legislators are ahead of those in the US, in part because social media is heavily used by terrorists;
- Some governments have moved to block known disinformation media streams in their countries, shielding their citizens from attempts at foreign influence;

- Many universities and private research groups have analysed disinformation campaigns, using distribution patterns and content indicators to identify bot networks and troll factories; and
- Specialised organisations have become skilled at exposing false news stories and, often in real time, educating the public to identify and expose disinformation.

Outlook

The negative impact on democracy of false news could increase if Russia and other actors become role models for others, increasing the distribution of malignant material through all the pathways of the electronic age.

Disinformation poisons public debate and is a threat to democracy. Raised public awareness is needed to distinguish the real from the false. There are many ways for governments and organisations to counter the threat, but there is no guarantee that even effective counter-campaigns can defeat the high volume flow of malicious communications.

CHAPTER 1

ORCHESTRATED OR EMERGENT? UNDERSTANDING ONLINE DISFORMATION AS A COMPLEX SYSTEM

Disinformation is spread through a complex network of often independent actors. Many are traffickers in conspiracy theories or hoaxes, unified by a suspicion of Western governments and mainstream media. Their narratives, which appeal to leftists hostile to globalism and military intervention and nationalists against immigration, are frequently infiltrated and shaped by state-controlled trolls and altered news items from agencies such as RT and Sputnik. Motivations for participation in the spread of disinformation are varied and should be taken into consideration.

Almost on a daily basis, new revelations expose the extent to which the Russian government used social media and other online tools to interfere with the democratic process in the United States, Britain and elsewhere. These discoveries illuminate a multi-dimensional strategy using high- and low-tech tactics to generate and spread disinformation. They also suggest a complex system in which these tactics resonate with and shape the activities of various types of distinct and independent actors.

Examining the spread of conspiracy theories surrounding terrorist attacks and mass shooting events in the United States can act as a lens for viewing the complex dynamics of this disinformation space. For example, after the Boston Marathon bombings, an online rumour claimed that the event had been a 'black ops' operation perpetrated by the US government. After the 2015 Umpqua school shooting, online communities of Reddit and Twitter users theorised that the event (like Sandy Hook three years earlier) was a 'hoax', staged by the government to justify gun control legislation. Similarly, the October 2017 shooting in Las Vegas was seen by some as a 'false flag' event carried out by members of the 'new world order'—a cabal of conspirators who pull the strings of world events.

These conspiracy theories are all somewhat distinct, but each reflects a pattern of claims about other man-made crisis events, and they all connect to a small number of shared underlying themes or narratives:

- The US government and other Western or NATO-affiliated governments are untrustworthy and are unjustified aggressors in conflicts around the world;
- These governments and other powerful people manipulate world events to ensure their power; and
- 'Mainstream' and corporate media are untrustworthy. They assist governments and other powerful actors in hiding the truth from people. They are 'fake news'.

Many of these narratives are explicitly connected to an 'anti-globalist' or nationalist worldview. The term globalism is a relative of globalization, used to characterise transnational perspectives[2] and policies that integrate free trade and open borders[3]. In practise, the anti-globalist term pulls people from seemingly disparate parts of the political spectrum onto common ground. For example, they connect left-leaning individuals who oppose globalisation and foreign military intervention by the US and other NATO governments with right-leaning individuals who oppose immigration and favour nationalist policies.

Tracking the spread of these conspiracy theories and their related narratives demonstrates how state-sponsored information operations interact with organic communities of online users to spread disinformation.

For example, on 5 November 2017, a mass shooting at a church in small-town Texas took the lives of more than 20 people. Within hours, officials and mainstream media identified a suspect, a 26-year-old man who had a record of domestic violence and had been discharged from the US Air Force. However, before that narrative developed, and then continuing even after it had been established, an alternative narrative claimed that the suspect was really an Antifa terrorist[4]. With the goal of forwarding this narrative, online activists on the political

right doctored screenshots of the shooter's Facebook profile to include an Antifa flag, providing evidence for this theory, and then used social media to spread that content. The theory soon began to propagate through the Twittersphere among alt-right accounts. Popular alt-right blogger Mike Cernovich tweeted that details of the shooter were consistent with the profile of an Antifa member. Alex Jones, a right-wing media personality known for spreading conspiracy theories, commented that the shooter wore all black (reflective of leftist activists). The theory also took root in alternative media, appearing on web sites like TheGatewayPundit, YourNewsWire and BeforeItsNews. Russian-government funded news outlet RT (formerly Russia Today) also helped to spread the claim, sharing a Facebook post that noted the shooter's Antifa connections, including content from the doctored Facebook profile.

> *State-sponsored information operations interact with organic communities of online users to spread disinformation.*

This activity follows a now established pattern of online activity after mass shooting events. Recent research suggests that some of the initial conversations around these theories take place in the less visible (and more anonymous) places of the Internet, such as Reddit, 4chan, Discord and others[5]. These theories are then spread and amplified, sometimes strategically, on Twitter and Facebook. Additionally, there exists a surrounding ecosystem of online web sites that takes shape around and supports these conspiracy theory-building conversations with additional speculation, discussion and various forms of evidence[6]. This ecosystem consists largely of alternative media that position themselves as challenging mainstream narratives. It includes several web sites and blogs that push conspiracy theories and pseudo-science claims (eg, InfoWars, 21stCenturyWire and SecretsOfTheFed). Significantly, many web sites in this ecosystem are news aggregators, remixing and republishing content found elsewhere in the ecosystem (eg, BeforeItsNews and YourNewsWire). For alternative narratives about shooting events in 2016, the system contains a few explicitly nationalist and white supremacist web sites

(DailyStormer) as well as some seemingly left-leaning activist web sites (ActivistPost). Web sites from the Russian-funded media outlets RT and Sputnik are also integrated into this ecosystem. Iran's PressTV appears as well.

An open question is how the different pieces of this dynamic system—of seeding, amplifying and spreading these theories—fit together. It is not yet clear how much of this activity is emergent and how much is orchestrated (and by whom and why). However there appear to be distinct actors, driven by varied and overlapping motivations. Six categories of motivation are proposed as part of a preliminary conceptual framework.

Sincere ideology. One set of actors within this system is ideologically motivated. These persons, including individual social media users as well as small organisations that operate web sites, blogs, and other feeds, are 'true believers' of the messages that they are spreading. The messages are largely anti-globalist (ie, anti-imperialism and anti-globalisation on the left; pro-nationalism and anti-immigration on the right). They are also explicitly critical and distrusting of mainstream media. These actors may indeed be affected by political propaganda, though causation is difficult to establish. At times, they can be seen to act as amplifiers of political propaganda, seeded with messages that they repeat and amplify. But many sincerely ideologically motivated actors also can be seen to generate their own content, without the continued need for direct seeding or coordination of messages.

> *...there appear to be distinct actors, driven by varied and overlapping motivations.*

Political propaganda. The activities of the second group of actors in this system, which include the intentional production, sharing and amplification of disinformation, can be viewed as part of a political strategy. Unlike the ideologically-motivated actors, these actors are not necessarily true believers of the messages that they share. In their messaging, they mix false information with factual information, and intentionally connect other stories and narratives, often the ones

that appeal to the ideologically motivated actors, to their own narratives. These politically-motivated actors are adapting old strategies of disinformation to the potential of the information age, leveraging the technological infrastructure of the Internet to spread their messages further, faster and at lower cost than ever before. Pomerantsev and Weiss[7] have written that the purpose of disinformation is not necessarily to convince, but to confuse—to create muddled thinking across society, to sow distrust in information and information providers. There is evidence that this strategy is at work within this system. Another goal of disinformation is to create and amplify division in (adversarial) democracies, and this is visible as well.

Financial incentives. Other actors within this system are financially motivated. For example, there are numerous web sites selling online advertisements and health products. Many are essentially aggregators of 'alternative' and 'pseudo' media, regurgitating 'clickbait' content designed to attract users. Others, like InfoWars, integrate original content with borrowed content from other sites in the ecosystem, including RT, and use their platform to peddle an array of products (ie, nutritional supplements).

Reputation gains. Another set of actors, particularly within the social media sphere, appear to be motivated specifically by the reputational and attentional benefits inherent to those platforms. Social media is designed to be engaging, and part of that engagement involves a feedback loop of *likes* and *follows*. In the disinformation space, especially among the alt-right, there appear to exist a set of actors who are primarily (or at least significantly) motivated by attentional and perceived reputational gains. Mike Cernovich and Jack Posobiec are two high-profile examples, but there are many others among the 'crowdsourced elite' on Twitter and elsewhere who spread alternative narratives and other politicised disinformation and have consequently received much online visibility.

The last two categories are more conceptual. While not yet backed by large volumes of empirical evidence, they are however theorised as part of this complex system.

Entertainment. It is likely that some participants in the disinformation space simply engage for entertainment value or 'for the Lulz', as the now waning Anonymous group would say. That slogan was meant to describe a kind of mischievous entertainment unique to online activity. Another way to think of this category is as extending gaming practices to the real world. For example, disinformation can provide a platform for working together with online team mates and an avenue for embarking on culture-hacking quests (to spread certain ideologies).

Empowerment. Disinformation can provide an opportunity for a disempowered individual or group to assert agency and power in the world through digital action. This category includes 4chan denizens who use memetic warfare[8] —the generation and propagation of graphical memes—to affect political change across the globe. Like digital volunteers who feel empowered by coming together online after disaster events in order to assist individuals, this set of actors is motivated by collectively working in an online team for a cause (eg, electing a favoured candidate). They are perhaps less motivated by the cause itself than by the emotional reward of having an impact.

These latter motivations and the associated sets of actors are significant. Preliminary research suggests that purposeful disinformation strategies are not just leveraging the power of social media platforms, but are resonating with the activities of online crowds that form within those platforms. For example, Russia-based troll accounts impersonating US citizens infiltrated online communities of alt-right Twitter users and functioned to both seed and amplify their messages during the 2016 US election cycle. They also embedded themselves within left-leaning Twitter communities that formed around issues such as #BlackLivesMatter, functioning to amplify existing divisions in the United States. On another front, Russia-connected information operations have targeted online activist communities that take shape around anti-war ideologies and use them to spread messages challenging US and NATO activities in Syria.

By focusing on explicit coordination by and collusion with state actors, and ignoring or under-appreciating the roles and motivations of these independent actors, researchers, journalists, and policy-

makers risk over-simplifying the complexity of this system, limiting the development of effective solutions, and under-informing public awareness of the problem. Importantly, the opportunity to assist everyday users of these systems to recognise the role they play within the disinformation phenomenon is missed. In other words, the problem of disinformation cannot simply be attributed to the design of technological systems or the deliberate actions of government-funded trolls. Solutions to this problem must also take into account the people who are interacting with and affected by this information, not merely as victims, but as agents in its creation, propagation, and (hopefully) its correction.

CHAPTER 2

RUSSIA, THE WEST AND THE GEOPOLITICS OF DISINFORMATION

The disinformation campaign carried out by the Kremlin and its connected oligarchical networks is a direct descendent of the KGB's 'active measures', increased in volume, speed and potency by modern technology. Its purpose is to control public opinion in Russia, and undermine Western democracies by creating division within targeted groups. Widely dispersed web sites, troll centres and hackers partly obscure the common origin of the fak and distorted news.

A century and a half before KGB Director Yuri Andropov made disinformation a central element of Soviet intelligence activity,[9] William Blake noted "A Truth that's told with bad intent Beats all the Lies you can invent"[10]. Such kernels of truth told with bad intent will be found at the heart of all disinformation properly defined, and are part of what makes disinformation so difficult to combat.

In this discussion, the adversary will be described wherever possible as 'the Kremlin' or other terms related to Vladimir Putin and his associates, rather than as 'the Russians' or 'Russia'. No good interest is served by representing the Kremlin's activities as Russia versus the West. In fact, the Kremlin's main adversary has always been, and still is, Russia itself. Virtually every type of action it has undertaken against the West was first implemented in Russia, against the Russian people, and against Russia's many ethnic, national and religious minorities. The Kremlin is a reference both to the presidential administration and the social networks of business leaders, organised crime bosses, as well as veteran officers, agents and assets of Soviet intelligence services, all of whom have ties to the Kremlin, and to Putin and his closest associates. This state-within-a-state, interacting with but standing apart from formal elements of the Government of the Russian Federation, has been described as an adhocracy[11]. People

move in and out of the presidential administration, performing tasks as needed, by turns acquiring or shedding what cover—or aura of legitimacy—a direct association with the Russian state may offer.

Disinformation, regardless of the entity engaging in the activity, is aggressive marketing of information in support of political objectives. The segmentation, targeting and positioning (STP) model has been a staple of marketing research and practise since at least the 1970s.[12] Social media platforms dramatically increase the amount of information available to guide the identification of market segments and the development of content most likely to influence the target audience. What is new is not so much the techniques, but rather the ease and rapidity with which disinformation can be simultaneously aimed at highly-segmented groups of people throughout the entire population of a country, at very little expense, and with little or no oversight or government regulation. Another important factor is the naïveté of technology companies, futurists, the general public and public policy-makers, who struggle to appreciate how much damage can be done to Western democracies by an unscrupulous adversary.

> *...the Kremlin's main adversary has always been, and still is, Russia itself.*

The methodology of disinformation may largely resemble contemporary marketing practise, but the stuff of disinformation, the content at the heart of the activity, is shaped by the political objectives being pursued, and by the absence of any moral or ethical constraints. Andropov himself defined disinformation by its observable effects, noting "Disinformation is like cocaine—sniff once or twice, it may not change your life. If you use it every day, though, it will make you an addict—a different man."[13]

We do not know if Andropov meant to suggest a physiological component to disinformation and its ability to capture the attention and compromise the mental capacity of those who consume it, but this may be a factor worthy of study. It is as though there is a 'disinformation receptor' in the human brain, and once stimulated,

this receptor convinces the brain that it must have more. The apparent physiological component of *disinformation* is likely enhanced by the many (largely negative) effects of computer-mediated communications and experience. The history of Soviet use of the term disinformation is itself an example of disinformation. First coined in Russia, the intelligence services of the Soviet Union and their allies were ordered in the early 1950s to spread a story indicating that the term was actually French, and described a weapon of information warfare deployed by the capitalist West against the USSR and people's democracies throughout the world.[14]

The Kremlin very much remains an adversary of the West. Putin and his associates are Andropov's children, recruited into the KGB in the 1970s as part of the Andropov levy, an effort to bring fresh blood and new ideas to bear on the many problems that beset the Soviet state.[15] While information technology in general, and the World Wide Web in particular, create new opportunities for the practise of disinformation, the playbook is largely unchanged. Just as jazz standards remain recognisable regardless of the players and the arrangements, so too do disinformation campaigns. By the time the Soviet Union collapsed, Western intelligence services had amassed an impressive body of knowledge regarding disinformation, and the larger set of tactics known as 'active measures'. Subsequent defections to the West and declassification of formerly secret reports mean we enter this new stage of antagonism with a much-improved understanding of what the Kremlin is doing, how, and to what ends.

Active measures had as their objective not intelligence collection but subversion. They sought to weaken Western countries internally, foster divisions among countries in the West, among NATO members and neutral European states, among the developed countries of Europe and North America and the developing countries of Asia, Africa and Latin America.[16] Soviet active measures targeted political leaders, opinion-makers, the media, business leaders and the general public of Western countries. The methods used went well beyond merely marketing information or promoting Communist ideology. False and deliberately misleading information was placed in the media; stolen and/or forged documents were leaked through cut-outs;

disruptive political movements were promoted where they existed and created where they did not; and subject matter experts were cultivated to shape policy in ways that served the Kremlin's interests. Aggressive use was made of diplomatic, commercial, academic and journalistic cover. Just as disinformation cannot be viewed apart from active measures, active measures are an integral part of Kremlin statecraft.[17] As it was then, so it is now.

But whereas before the West was confronted by a monolithic Soviet state, today's Kremlin adhocracy provides new opportunities for combatting its efforts. While much attention has been paid to a single Kremlin troll factory in Saint Petersburg, the fact is much of what can be observed with regards to disinformation and other active measures online is as likely to originate from an advertising agency in Zurich, for example. Acting at the behest of current officers of Russian military intelligence (GRU) in Moscow, a 'Patriotic Journalism' club in Omsk, in Russia, may create an alternative media web site purporting to cover conflicts in the Middle East. The women in Omsk, who answer to a board of directors composed of veteran Soviet GRU Spetsnaz officers, make use of services provided by ethnic Russian criminal hackers in Spain, who have servers in a data centre in Amsterdam and an address of convenience in Hong Kong. All this to bring a web site online for a team recruited from among retired analysts formerly employed by Warsaw Pact intelligence services. This scenario is not uncommon, and while tracing the lines of communication back to Moscow may take time, the nature of the personnel involved in the operation means tradecraft will be inconsistent and oftentimes ad hoc, creating investigative opportunities.[18]

> *The attraction of disinformation appears directly associated with the attraction of authoritarianism.*

What is to be done? Disinformation can be confronted on many levels. The most pernicious effects can be mitigated. Targeted populations can be rendered more resistant. Both non-state and state-sponsored purveyors can be confronted, convinced—one way

or another—to cease and desist. To the extent human populations are hard-wired to accept disinformation, to believe the worst of their fellow humans, there will never be a total victory. The attraction of disinformation appears directly associated with the attraction of authoritarianism. Democratic Western pluralism is vulnerable for the same reasons it is valuable. Without effort, it will not survive. Certain truths need to be inculcated in each generation, first among them that there is such a thing as truth—that there is an objective reality that cannot be wished away. There is a need to understand how technology exacerbates the problem of disinformation, and if possible find ways to alter how information is delivered in order to affect how it is received and experienced by each of us. Enemies both foreign and domestic who use disinformation to undermine democracy and the rule of law must be confronted and exposed for what they are: subversives. It has taken centuries of concerted effort to raise societies above humankind's more base, destructive and intolerant tendencies. Finally, those who are involved in the study of disinformation, who publicly confront the issue, and the state and non-state actors engaged in the activity, need to keep in mind that there are no passive observers. There are no front lines—the war is total—and there is no neutrality. Driving wedges between people is sure to be one objective of the Kremlin, and it is incumbent upon everyone to make an effort to not be pawns in a Kremlin game.

CHAPTER 3

NATO'S EASTERN FLANK: A NEW BATTLEGROUND

Russia has developed a comprehensive information strategy with the objective of increasing its influence over the Baltic-to-Black Sea periphery, while increasing its potential for successful military action in any future confrontation with the countries on NATO's eastern flank. Spreading non-ideological and targeted information is aimed at diminishing the will of targeted populations to resist Russian dominance, while discrediting NATO forces pledged to come to their assistance.

Among other crucial developments, the year 2013 witnessed Russia openly declaring an information war on the West. The first wave of onslaught was directed against states placed between the Baltic and the Black Sea (so-called 'Intermarium'): Estonia, Latvia, Lithuania and Ukraine—countries that have remained the prime target of Russian intimidation and aggressive behaviour since the dissolution of the USSR in 1991. Aside from this, Kaliningrad Oblast, an exclave of Russia, emerged as a unique case study illustrating Russia's resolve to build an anti-Western 'ideological bastion' in the heart of Europe. Given critical role of information in Russia's vision of the future of warfare, the campaign against Ukraine and the Baltic states is nothing else but an integral part of the Kremlin's general preparation for future conflicts.

Making a Molotov cocktail: Information warfare à la russe

Russia's current disinformation campaign against the West is more dangerous and sophisticated than ever before for several reasons. First, the Soviet strategy wrapped in modern attire makes it universal, flexible, smart and borderless. Second, hacking campaigns, *kompromat* attempts, the deliberate destruction of information, blatant corruption, and cyberattacks render it virtually untraceable. Third, designed for domestic and foreign consumption, it reaches out to different

audiences in Russia, the post-Soviet space and beyond. Fourth, it is permanent: known as 'informational confrontation', it is designed for both war and peace time. Finally, it often contains seeds of truth, which makes it even more difficult to defeat.

> *Russia's current disinformation campaign against the West is more dangerous and sophisticated than ever before for several reasons.*

Russian disinformation is extremely flexible. While the West is struggling to fit it in any theoretical framework, the Russian side is merging theory and practise as part of a multi-disciplinary approach, weaponising information. Thus, a combination of Soviet-inspired post-modern information-psychological and information-technology warfare constitutes two parts of the same phenomenon.

Following the outbreak of the Ukrainian crisis in 2014, Russian disinformation efforts became more sophisticated and gained some new features, as described below.

Militarisation of information. Russian military strategists consider disinformation as an organic part of future conflict. Theoretical research and practical steps resulted in the creation of 'research units' and 'cyber troops'. According to Russian Minister of Defence Sergey Shoigu, those "will be much more efficient than the 'counter-propaganda' department of the Soviet period".

Codification and renovation of information legislature. The adoption of a new information doctrine (2016) and strategy for the development of an information society (2017) has tightened the state's control over the domestic information space, identified external priorities, and confirmed Russia's readiness for information warfare.

Creation of the 'information vertical'. Every Russian citizen, from the President to a local operator, is now a part of centralised

vertical responsible for the state's information security. Introduction of 'cyber squads' and the extension of the Russian National Guard's responsibilities in the domain of information and cyber security is part of this strategy.

Ukraine: Russia's laboratory for future wars

The Russian disinformation assault against Ukraine corroborates Lenin´s tenet that 'propaganda should be a matter of action rather than words'. The post-2013 developments should be viewed as a logical conclusion of the Kremlin's previous sustained covert actions since the early 1990s. Russia's invasion of Ukraine witnessed the combination of kinetic and non-kinetic methods simulating a new type of military conflict, with local military action (conducted by special-operations forces) supported by disinformation campaigns and cyberattacks. The first stage, the annexation of Crimea, served as a springboard for subsequent events in the Donbas region. The Russian side employed both information-technology (the occupation of the Simferopol Internet Exchange Point and disruption of cable connections to the mainland that secured Russian information dominance over the peninsula) and information-psychological warfare targeting Ukraine and the EU. At this juncture, emphasis was given to reflexive control techniques, when Moscow attempted to force the international community to recognise Russia as an actor with special vested interests in Ukraine, while at the same time supposedly not being a party to the conflict. The second stage of the conflict, from April 2014, saw a similar but expanded strategy based on intensified disinformation efforts, cyberattacks, troll farms and botnets, IT software and search engines (primarily Yandex) as a means to defeat, discredit and falsify information. Russia's attempts to discredit Ukraine in the eyes of the West were based on presenting it as a 'mistake of 1991', a failed stated ruled by illegitimate, corrupt, inefficient, Russophobic, anti-Semite neo-Nazi 'junta'—arguments that were to reach out to every segment within Western society.

The ruthlessness and actions of Moscow hinged on the following assumptions:

1. Moscow would not be challenged over Ukraine;

2. Weak, disunited and lacking strategic vision, Ukrainian political elites would fail to react properly; and

3. Ukraine is not a (homogenous) state, meaning that Russian actions will be supported in certain regions.

Worst of all, for the majority of Ukrainians the idea of war with Russia was inconceivable, which was cynically abused by the Kremlin. In this regard, the Ukrainian example should be recognised as a stern warning to the entire European community, and to the Baltic states in particular.

The Baltic states: The next targets?

The three Baltic states comprising the northern part of NATO's eastern flank are another prime target of Russian disinformation. Throughout the 1990s, Russian propaganda efforts revolved around the interpretation of Soviet historical legacy, with many poorly integrated and Soviet-nostalgic Russian-speaking minorities acting as the Kremlin's 'fan club'. After 2007, dramatic changes owing to the emergence of the 'Russian world' concept ensued: the once poorly organised and frequently incoherent actions of the Russians evolved into a systematised, well-coordinated and coherent strategy.

Russia's disinformation operations against the Baltic states aim to present these countries as a failed experiment of both post-Soviet transformation and Euro-Atlantic integration. Russian propaganda extensively draws on 'widespread poverty', 'depopulation', raging far-right ideology and the 'semi-colonial status' of these countries. Meanwhile, the local elites are portrayed as Russophobic and paranoid. According to Russian propaganda, these features, coupled with 'blind servility' to the West, do not allow local political elites to make rational decisions, damaging their economies and turning these countries into a 'sanitary cordon' against Russia, and at the same time a target for Russian retaliation.

Another crucial theme of Russian disinformation is inseparable from the role of NATO. Kremlin-backed propagandist outlets are spreading fake materials and news (in Russian and local languages), attempting to create a repulsive image of NATO, whose soldiers (especially in Lithuania and Latvia) are portrayed as a wild mob of vandals, sexual perverts, and rapists immune to local laws and acting like invaders (an apparent parallel with the Nazi army on the Soviet territory). This distorted narrative serves the following objectives:

- Internal mobilisation of Russian population around the current political regime ('Russia as a besieged fortress');
- Russia as an alternative to the Western-liberal model ('Russia as custodian of Christian-conservative values');
- Revival of anti-American/NATO sentiments in Europe; and
- Artificial fragmentation of the EU.

Another way to create a negative image of NATO relates to the massive military build-up in the Western Military District (in particular, Kaliningrad Oblast), which aims to create an aura of impunity and at the same time 'prove' to the Baltic states that NATO is powerless to protect their sovereignty and territorial integrity in the event of conflict. At the same time, Russian military escalation attempts to stress the point that 'excessive' military expenditures are nothing but an unnecessary waste of money (and NATO-imposed condition) that could have been invested in the economy instead.

The Ukrainian crisis has had a dramatic impact on Russia's behaviour in regard to Estonia, Latvia and Lithuania. The most recent aggressive actions have portrayed the Baltic states as nothing but a 'near abroad', entities that have not escaped the Russian sphere of interest while at the same time failing to join the Euro-Atlantic community. Aggressive disinformation campaigning against the Baltic states is also meant to show that growing tensions in the region are caused by anti-Russian actions and Russophobia spreading in the Baltic states and Poland, which according to senior Russian officials could cause the Third World War.

Sabre rattling and direct intimidations are merely one side of Russia's changing posture. After 2014, by spreading fake materials and aggressively interfering in the domestic affairs of its neighbours, Moscow has been increasingly leveraging Kaliningrad as a new outlet, while generating anti-Polish, anti-Lithuanian and anti-NATO sentiments.

Kaliningrad: Beacon of the 'Russian world' in Europe.

Russia's ability to act in the Baltic states and Ukraine is constrained by a number of factors and is likely to be limited to an even greater extent given realities of the post-Crimean world. Located in the heart of the EU, Kaliningrad appears to be an ideal location for the generation of disinformation and export of Russian values abroad. First attempts to that effect were unsuccessfully made from 2003 to 2006. However, it was the Ukrainian crisis that became a genuine game-changer, transforming the Kremlin's perception of Kaliningrad, and its role in the ideological conflict with the West.

From 2014 on, the exclave has been in the vanguard of vigorous anti-Lithuanian, anti-Polish disinformation campaigns. The most notorious example was a disgraceful episode in Vilnius at the end of 2016, when the Russian embassy disseminated propaganda leaflets with fraudulent data on Lithuanian economic performance, urging the locals to abandon the country for Kaliningrad.

> *Kaliningrad has become a shield of the so-called Russian world in an ideological war against the West.*

Apart from stoking internal disturbances Kaliningrad has become a shield of the so-called Russian world in an ideological war against the West, its values and traditions, a world in which the Russian Orthodox Church (ROC) has acquired prominence. Speaking in Kaliningrad (March 2015), Russian Patriarch Kirill named the oblast "Russia's beacon" and a shield against the "adverse world". Coupled with breath-taking militarisation (resulting in the oblast becoming one of the most formidable anti-access/area-denial regions), Russia's

measures in the domain of information security have transformed Kaliningrad into a laboratory for testing future warfare, with both sides of Moscow's information confrontation being used in an integrated strategy.

What comes next?

From the Black to the Baltic Seas, NATO's eastern flank presents a relatively weak, fragmented and unevenly developed area. Given the lessons Russia has drawn from its experience in Syria and Ukraine, Moscow will stress pursuing a strategy based on an integrated use of military and non-military components. As described by Chief of the General Staff Valery Gerasimov (2016) "the emphasis on the method of fighting [is moving] toward[s] the complex application of political, economic, information and other non-military means, conducted with the support of military force". This means that the notion of information security should be seen as an organic part of hybrid warfare.

Furthermore, there is every reason to believe that another Russian strategic objective is concerned with undermining the level of cohesion among EU and NATO member states, as well as generating conflict and mutual animosity between Ukraine and its strategic partners in the Euro-Atlantic alliance. This will be done using various means from Moscow-backed think-tanks, NGO's and marginal populist groups to social media and information outlets. The sophistication of Russian propaganda requires the West to abandon what has often been a simplistic understanding of information warfare.

CHAPTER 4

FOREIGN INFLUENCE EFFORTS AND THE EVOLUTION OF ELECTION TAMPERING

After successes in the Arab Spring and the Russian election of 2011-12, the Kremlin increased its use of information operations and *kompromat*. Many techniques are employed to make disinformation appear genuine, including selecting television interviewees who will provide a pro-Moscow interpretation of events on state-controlled channels and exploiting both human and automated dissemination techniques to distribute faked stories to those willing to mount dissent within foreign political systems.

The central concept to understanding Russian information-influence operations beyond the country's borders is the 'protest potential of the population'. This term is included in Russian military doctrine[19] as one of the main features of modern (not just Russian) conflict, alongside military activities, political, economic and informational tools, as well as special forces. The term was introduced in the doctrine after the events of the Arab uprising of 2011, and the widespread protests against vote-rigging in Russia in 2011 and 2012. According to Russian official statements, Western powers staged these protests to topple pro-Russian regimes.

The Kremlin's initial reaction was to target Russians, to prevent any recurrence of democratic enthusiasm. Initiatives such as the 'foreign agent's law', cracking down on pro-transparency NGOs, stem from this period. Simultaneously, a troll factory—Russians paid to make political posts online—was established in St. Petersburg to flood Russian opposition communities with pro-government posts. Russia served as a test-bed for these methods; the government's first goal, as so often, was to ensure its own survival. Subsequently, and especially after the Crimean annexation in 2014, the same weapons were extended to international targets, first to Ukraine, then to the West.

Approach

Russia's approach to information-influence operations in democratic states can be summarised as 'vilify and amplify'. Different parts of the Kremlin's systems generate or gather material designed to undermine the target; the other parts of the system amplify that material, while preserving a degree of plausible deniability. This method dates back to pre-Soviet times and the concept of *kompromat* (from 'compromising material'). In the 1980s, the Soviets posted a fake claim in an Indian newspaper that the CIA had created AIDS, and then amplified it worldwide. The advent of deniable web sites and social media has made such techniques much easier to deploy.

One simple technique is to give a platform to commentators in the target country who validate the Kremlin's narrative. For example, in 2014 and 2015, RT interviewed a disproportionately high number of members of the European Parliament from Britain's anti-EU UK Independence Party (UKIP); in the first half of 2017, Sputnik France devoted disproportionate coverage to politicians who attacked Emmanuel Macron. During the US election, RT and Sputnik repeatedly interviewed an academic who claimed that Google was rigging its auto-complete search suggestions to favour Clinton.

In such cases, what is important is what is left out, as much as what is included. The interviewees can be, and usually are, sincere in their beliefs; the propaganda technique consists of amplifying and validating those beliefs without providing the other side of the story. RT has repeatedly been found guilty by the UK telecommunications regulator in this regard.

> *What is important is what is left out, as much as what is included.*

Close analysis of the 'experts' themselves is also important. For example, in the build-up to the Catalan referendum on 1 October 2017, Sputnik's Spanish service headlined tweets from Wikileaks founder Julian Assange more than any other commentator, including the Catalan president or Spanish prime minister. Assange had never

mentioned Catalonia in tweets until 9 September 2017; he is not known to have any special expertise in Spanish constitutional affairs. Sputnik's decision to amplify his tweets, which attacked the Spanish government, therefore appears based on his message, rather than any expertise.

Fake experts: Partisan commentators

A separate technique is to plant comments from Kremlin-aligned speakers without mentioning their affiliation. For example, after the shooting-down of Malaysian Airlines flight MH17 over Ukraine, investigative journalists with the Bellingcat group gathered evidence from open sources demonstrating that the plane was shot down with a Buk-M1 missile which had entered Ukraine from Russia.

In response, a group of initially anonymous and 'independent' bloggers calling themselves 'anti-Bellingcat' published a lengthy report rebutting Bellingcat's findings. The anti-Bellingcat report was widely reported in multiple languages by Kremlin outlets.

It later emerged that, far from being independent, one of the two lead authors worked at the state-owned company which produces the Buk missile; the other was spokesman for a Kremlin-founded think tank linked to Russian intelligence.

Kremlin bodies also have created a number of 'independent' sites which mask their ties to the Russian government. NewsFront.info, for example, produces pro-Kremlin and anti-Western content in a number of languages; according to a whistleblower interviewed by *Die Zeit*, it is funded by Russian intelligence. A collection of web sites in the Baltic states, Baltnews, claim to be independent, but have been traced back to Sputnik's parent company. In October 2017, a highly active and influential far-right US Twitter account, @TEN_GOP, was exposed as being run from the troll factory. This account was extraordinarily successful—quoted in the mainstream media and retweeted by key Trump aides—amplifying disinformation which was eventually quoted by Trump himself.

The same month, a group known as AgitPolk ('agitation regiment') was outed as being tied to the troll factory. This group posed as online activists, and repeatedly launched pro-Kremlin or anti-Western hashtag campaigns, including attacking US actor Morgan Freeman and wishing Russian President Vladimir Putin a happy birthday. On one occasion, unknown actors created a complete mirror web site of *The Guardian* to post a story claiming that the former head of MI6 had admitted that the UK and US had tried to break up Russia in the early 2000s. The fake was quickly exposed, but this did not stop Russian state TV from running lengthy reports on the story, validating their narrative of a Russia under siege.

The most damaging technique is hacking the emails of target politicians, and leaking them online. This is especially harmful because:

- there is an implicit assumption that any leak must be damaging;
- it is easy to insert faked documents amidst the real ones;
- leaks can be held back until the most damaging moment; and
- in an unsuspecting environment, real media are likely to amplify the leaks.

The hacking of emails from the campaign of US Democratic candidate Hilary Clinton, and their leaking online, fits squarely into this *kompromat* pattern. The leaks were used particularly aggressively, with a selection being published daily in the month before voting day. The intent of these operations appears to have been two-fold: to undermine Clinton personally, and to attack the legitimacy of the election process in general. This was done in the hope of galvanising the 'protest potential of the population' in the event of a Clinton victory. It is one of the ironies of 2016 that Clinton lost, and that Russia's interference in fact undermined the president it had boosted.

Another divisive technique which is still being exposed is the practise of buying partisan advertisements for placement on social media. Combined with the use of anonymous and aggressive social-media

accounts, this technique appears designed to pit multiple groups with protest potential against one another.

Developments

Given the widespread exposure of recent techniques, we can expect them to evolve rapidly. Adaptations are likely to aim at masking attribution more effectively, and blurring the distinction between human and automated operators. We have already seen efforts to reduce the danger of leaks from the troll factory through a heightened insistence on patriotism among staff[20]. It is also noteworthy that, while the Clinton campaign emails were leaked via Wikileaks, emails hacked from Macron's campaign were dumped anonymously on 4chan, a web site, and amplified by the far right in the US, suggesting a desire to vary the delivery platform.

Social-media accounts are becoming increasingly sophisticated in their combination of human-authored and automated posts. Such cyborgs typically post at high rates, in the hundreds per day, but intersperse these with authored posts, making them less obvious to bot-detection algorithms, and harder to counter. This trend is likely to accelerate.

Hacking attempts can be expected to grow, especially from deniable actors whose links to the Kremlin are masked. The experience of 2016 showed that hacking and leaking can be a devastating weapon, but that this can backfire if the hacks are attributed. It is likely that the leaks attacking Emmanuel Macron were published anonymously on 4chan and spread by the far right in the US in an effort to make attribution still more difficult. A move away from overtly Kremlin-owned outlets such as RT and Sputnik may also materialise, as these come under increasing scrutiny, with a greater emphasis on front outlets such as NewsFront and the BaltNews family.

Countermeasures: Building resilience

A number of disinformation countermeasures have already been trialed. The simplest has been to block the accreditation of pseudo-

journalism outlets such as RT and Sputnik, as was seen in the Baltic states and France. This approach sends a powerful signal, but also sets a precedent which can be open to abuse. Such moves should only be used as a last resort.

> *Hacking attempts can be expected to grow, especially from deniable actors whose links to the Kremlin are masked.*

Registration of state-controlled media is also an avenue worth pursuing; at the time of writing, RT and Sputnik are reportedly facing demands to register as foreign agents in the US. Again, such approaches must be measured: the key is to label the outlet without giving the impression of silencing it.

Regulation of journalistic standards can also play a part. In the UK, the national telecoms regulator, Ofcom, has found RT guilty of breaching journalistic standards in a number of broadcasts. The sanctions have been symbolic; the reputational damage has been considerable. Such regulatory findings, based on the detail of individual programs, and pegged to transparently-defined standards of due accuracy and impartiality, are a valuable tool in efforts against all disinformation, from all sources.

Detailed fact-checking also has a part to play in debunking false stories and narratives. Given the emotional nature of most fake stories, fact-checking is not best suited to countering a specific story; however, over time, a regular pulse of fact-checking can help to expose key sources of fakes. Exposing influence attempts is also important. In the best case, such as recent fake allegations of rape against NATO soldiers in the Baltic states, rapid official engagement with the mainstream media to expose the attempt materially contributed to those stories' failure to gain traction[21].

However, for such exposure to succeed, there must be a degree of understanding in the media and in society that influence operations are dangerous, should be taken seriously, and should be addressed

promptly. Brushing aside the issue can have consequences. The US Director of National Intelligence warned, on 7 October 2016, that Russia was attempting to interfere in the election. Quickly drowned out by the release of the Access Hollywood tapes in which Trump boasts about grabbing female genitalia, the warning only gained nationwide traction after the election.

The importance of education and engagement with the population cannot be overstated. Disinformation spreads best in groups which are unsuspecting or who are biased in favour of the fake. Online literacy skills, such as how to identify a fake social media account, stolen photo or tendentious article, should be taught far more widely; governments might also invest more in identifying, engaging with, and listening to, particular segments of their societies, to understand how and why fake stories spread among them.

There is no single answer to the complex and multi-faceted nature of disinformation. Regulation, fact-checking, exposure and education all have a role to play; a response which highlights just one, while ignoring the others, can be expected to fail. The solution is to boost resilience on as broad a front as possible.

CHAPTER 5

EXAMINING BREXIT: THE RISE AND FALL OF A TWITTER BOTNET

Research on botnets operating during the Brexit referendum shows a pattern of coordinated hyper-partisan tweeting which featured one stream generating automated tweets and retweets in high volumes, and a second stream distributing user-generated material to a more targeted readership. A majority of traffic favoured the Leave side, and appealed to nationalistic and xenophobic readers. While not deliberately-constructed faked news, content was often fact-free and simplistic, mirroring the style of the tabloids, and incorporating reader feedback loops. A high proportion of the accounts, and their related content, were terminated immediately after the referendum.

The referendum on the UK's membership in the European Union was held against a backdrop of political realignment, polarisation, and hyperpartisanship. Additionally, news readership mirrored a demographic splintering, dividing news consumption along broadsheet and tabloid media outlets. Those elements were strategically leveraged and maximised by populist parties and leaders during the referendum in order to promote "traditional cultural values and emphasize nationalistic and xenophobia appeals, rejecting outsiders and upholding old-fashioned gender roles"[22]. These circumstances and the political climate which resulted offered fertile ground for bot activity during the Brexit referendum.

The following analysis examines the activity of a botnet that tweeted the referendum by sourcing a range of user-generated and user-curated content featuring hyperpartisan reports. Thirty-nine Twitter hashtags clearly associated with the referendum campaign from April to August 2016[23] were analysed, which collectively amounted to 10 million tweets. Subsequently, the profiles of over 800,000 unique users were retrieved, and thresholding and filtering approaches were implemented

to disentangle real users from bots. A combination of methods were used to identify a large group of bots whose accounts had been deactivated by the bot master or blocked/removed by Twitter in the aftermath of the referendum; identify the campaign associated with the tweets; retrieve the web page title of URLs embedded in tweets (when available); and examine retweet and @-mention behaviour.

Disappearing tweeters

From a total of 794,949 Twitter profiles that tweeted the *Vote Leave* and *Vote Remain* campaigns, five per cent were identified to have been deactivated, removed, blocked, set to private, or to have altered their username after the referendum. Of this group, the majority (66 per cent) had changed their username since the referendum but remained active on Twitter (repurposed or recycled accounts), and 34 per cent were suddenly blocked or had removed themselves from Twitter (deleted accounts). Common among recycled and removed accounts is the predominance of retweeted content that disappeared from the Internet shortly after the referendum. Another commonality is the notable support for the *Vote Leave* campaign, measured by the relative frequency of keywords and hashtags associated with each of the campaigns. While the total ratio of messages using hashtags that supported the *Vote Leave* and *Vote Remain* campaigns was 31 per cent and 11 per cent respectively, recycled and removed accounts combined tweeted the referendum hashtags to a ratio of 37 per cent and 17 per cent.

Analysing the language of the tweets provided additional insight into this disparity. By annotating tweets using textual markers such as hashtags and keywords associated with the *Vote Leave* and *Vote Remain* campaigns, the proportion of tweets supporting the *Vote Leave* campaign in the pool of removed accounts was yet higher, at 41 per cent compared with 31 per cent for active users, with the proportion of neutral tweets also being higher in the latter. Slogans associated with the *Vote Leave* campaign were also significantly more likely to have been tweeted by this pool of accounts in a ratio of eight to one. This subset of removed accounts was considerably more active in

the period leading to the referendum, and also less active in the wake of the vote.

Hyperpartisan and hyperperishable news

Attempts to retrieve the web pages tweeted by recycled and removed accounts found that most tweeted URLs (55 per cent) no longer existed, could not be resolved, or linked to either a Twitter account or web page that no longer exists. Nearly one third (29 per cent) of the URLs link to Twitter statuses, pictures, or other multimedia content that is no longer available and whose original posting account has also been deleted or blocked, a marker of the perishable nature of digital content at the centre of political issues. Of this total, one per cent of all links was directed to user @brndstr, one of the few accounts appearing in the communication network of recycled accounts that remains active under the same username. This account is managed by a company which specialises in providing bots for social media campaigns.

A closer inspection of the accounts sourcing content to the pool of recycled and removed accounts reveals the markedly short shelf life of user-generated content. These are Twitter accounts invested in spreading dubious news stories sourced from a circuit of self-referencing *blews*: a combination of far-right weblogs and traditional tabloid media. However, the few retrieved web pages indicate that the content tweeted by this pool of recycled and removed accounts do not conform to the notion of disinformation or fake news. Instead, the content is in line with a form of storytelling that blurs the line between traditional tabloid journalism and user-generated content, which is often anonymous, fact-free, and with a strong emphasis on simplification and spectacularisation. User-generated content takes the lion's share of hyperlinks tweeted by recycled and removed accounts, often presented as a professional newspaper via content curation services, and is likely to include Twitter multimedia.

Similarly, the few links that remained accessible six months after the referendum consisted of material rich in rumours, unconfirmed events and human-interest stories with an emotional and populist

appeal that resembles tabloid journalism, with the added complexity that audiences play a pivotal role in curating and distributing the content. The inspected sources, though not representative of the much larger universe of content tweeted by this population of users (and which has unfortunately mostly vanished from Twitter), is much akin to hyperpartisan tabloid journalism, with a topical emphasis on highly-clickable, shareable and human-interest driven stories.

Although 17 per cent of weblinks pointed to Twitter accounts that are still active, an examination of a random sample shows that the original message is frequently no longer available, thus preventing any determination of the nature of the content originally tweeted. For example, one profile generated a cascade of several hundred retweets and was found to have an active posting user. Although the user account seeding the cascade remains active, the original tweet has been removed (together with the relevant retweet cascade). With Internet Archive having no record of this specific tweet, it is no longer possible to know what the original image conveyed. The scale of deleted content applies both to weblinks tweeted by this population as well as to user accounts, a worrying development given the importance and contentious nature of the referendum.

Brexit Botnet

Subsequent inspections surrounding the retweet behaviour of bots shed light on the existence of at least two clusters of fundamentally different bots. The first group was dedicated to replicating automated content, often hyperpartisan news, hence achieving a much faster cascade turnaround compared with active user-generated cascades. The second group was deeply embedded in human-driven activity. Both types of account succeeded at generating medium ($S>50$) and large cascades ($S>100$), but their typical retweeting patterns indicate they were created and deployed to meet fundamentally different objectives.

While the first subset of bots was associated with accounts that leveraged retweet behaviour to amplify the reach of a small set of users and rarely, if ever, started any cascade themselves, the other

subset of bots had a narrower scope of operation, only retweeting other bots in the botnet and thereby producing many medium-sized cascades that spread significantly faster than the remainder of the cascades. Although both are bots, the first only retweets active users, whereas the retweet activity of the latter is restricted to other bots (likely deployed in conjunction with the head node). Each of the bot subnets plays a specialised role in the network, and both feed into the larger pool of regular accounts brokering information to @vote_leave, the official Twitter account of the *Vote Leave* campaign, and arguably the most prominent point of information diffusion associated with the *Vote Leave*.

> *Inspections surrounding the retweet behaviour of bots shed light on the existence of at least two clusters of fundamentally different bots.*

Retweet activity was mostly concentrated in the period leading up to the referendum vote. Most of it consisted of organic retweets from and to accounts in the active user base. Bots operated in the same period both by retweeting active users and retweeting other bots, mainly in the week preceding the vote and on the eve of the referendum, when a peak in retweet activity between bots was observed. There was a sharp decline in retweet activity after the referendum, mainly among active users who ceased to trigger or join retweet cascades. Bots remained operational throughout the campaign and activity peaks were observed in the period from 12 to 15 July: first retweeting active users, then replicating bot content, only to tail off in the following weeks when the botnet was retired, deactivated, or removed entirely from the Twitter platform[24]. In fact, head nodes of the bot-to-bot subnet mostly disappeared after the referendum. This is the critical period when content tweeted by such bots and the web pages linked to their tweets disappeared from the Internet, Twitter public, and enterprise application programming interfaces (APIs).

Conclusions

The large number of links directed to user-generated content, particularly Twitter multimedia and the significant incidence of content curation services used to render socially shared content into professionally-looking online newspapers suggests that the universe of hyperpartisan news is both engineered top-down and reliant on user-generated content. While the content tweeted on Brexit has a stronger slant towards nationalist and nativist values compared to the content tweeted by the global population (27 per cent versus 19 per cent, respectively), the emerging reality of hyperpartisan web sites is that they cater to both extremes of the political spectrum, are often owned by the same companies, and repurpose stories to accommodate and confirm readership bias.

...the emerging reality of hyperpartisan web sites is that they cater to both extremes of the political spectrum, are often owned by the same companies, and repurpose stories to accommodate and confirm readership bias.

Analyses of the Brexit botnet did not find strong evidence of widespread 'fake news' dispersion, but rather surfaced the strategic placement of bots to feed user-curated, hyperpartisan information. The results presented in this study point to another milestone in tabloid journalism: the ability to incorporate an audience feedback loop while transitioning from the editorial identity of traditional tabloid newsprint to content curation that is both user-generated and created by editorial staff. Hyperpartisan news outlets thus epitomise the ongoing trend to churn out viral content that is mostly short, highly visual, shareable, accessed through mobile devices, and that, by confirming audience bias, sits side by side with the balkanisation of readership according to interests of like-minded groups.

CHAPTER 6

APPLYING OPEN-SOURCE METHODS TO DEBUNK 'FAKE NEWS' ABOUT SYRIA

Russia's military intervention in Syria has preserved the Assad regime while denying all accusations of illegal tactics and war crimes in the area. However, the Syrian case study illustrates that Russia's technology-driven weaponisation of information can be countered by that same technology. Open sources provide digital fragments that can be gathered and cross-referenced to disprove propaganda and provide direct evidence on Russian tactics.

From analogue to digital

Fake news, disinformation, propaganda, no matter the term, the challenge of disinformation has reached a new level of complexity in a hyperconnected world. The days in which information flowed in one direction, from governments, publishers and broadcasters to the public are over. Today, every smartphone user can be broadcaster as well as consumer, reporter as well as reader. This tectonic shift only began a decade ago, but already more than 3.8 billion people have access to the Internet; 2.9 billion are social media users; and 2.7 billion are mobile social media users.

This revolution presents potent new tools for the study of conflicts, crises and disinformation and has motivated an entire movement of so called Digital Sherlocks to focus on methods that help filter through the fog of disinformation. Conflict zones and hotspots that were once unreachable can now be accessed through online posts. Hostile disinformation actors are aware of the opportunities this new environment presents and are working around the clock to exploit this information and undermine the basic principles of reality.

Background of the Syrian conflict

The case of Russia's role in Syria underscores the challenges posed when a state actor utilises disinformation and deception to back its acts of aggression. Such methods allowed Russian President Vladimir Putin, in the last few years, to move from one foreign policy adventure to the next, in the process weaponising information against Western societies.

In 2014, Putin ordered the annexation of Ukraine's Crimea, overseeing a clandestine war in eastern Ukraine and backing Russian proxies with weapons, fighters and entire army units. As that war ground down into stalemate, Putin turned his eyes to Syria. After a rapid diplomatic campaign, and an equally rapid military build-up, he launched air strikes in the war-torn country. Russia's military campaign allowed Assad's forces to retake lost ground, a task they completed with great brutality and immense human suffering. Far from shortening the war, it exacerbated it, and in so doing, it sent yet more waves of refugees flooding into Turkey and Europe. None of this would have been possible without the veil of disinformation under which Putin and the Assad regime covered their actions and atrocities.

The veil

Putin cynically claimed that Russia's presence in Syria was aimed at fighting Daesh, openly encouraging the myth that Russia was fighting terrorism, that the Assad regime was innocent of atrocities, and that the Syrian uprising was instigated by the West. The veil was successfully held in place by employing three strategies:

1. *Denying the deed.* The simplest response to allegations of civilian casualties and indiscriminate strikes was to deny them. Throughout the conflict, and in defiance of the evidence, both the Syrian and Russian governments rejected such allegations outright.

2. *Militarising the victims*. In parallel to the campaign of denial, Syrian and Russian officials repeatedly misidentified their

targets, presenting civilians as combatants. This re-branding of civilians as legitimate military targets covered both entire city areas and individual buildings. By repeatedly blurring the distinction between Al-Qaeda-linked forces and other groups, Russia and Syria were able to create an impression that all groups targeted by them were extremists.

3. *Attacking the witnesses*. As became particularly clear during the siege of Aleppo in 2016, eyewitness evidence could discredit the Russian and Syrian attempts to militarise victims; airstrikes were hitting civilian buildings and civilians were dying. In response, Syrian and Russian officials began to attack the credibility of such witnesses. One of the most important witnesses to the suffering was the aid organisation initially called Syria Civil Defence, later dubbed the 'White Helmets' after its staff's trademark headgear. In Aleppo, the White Helmets began as a rescue organisation in early 2013[25]. As the conflict intensified and independent journalists no longer had access to the front lines, the White Helmets increasingly became a main source of evidence of the true nature of the bombings, posting GoPro footage of airstrikes and their aftermath. This put them on a collision course with the government and its allies.

Those seeking to spread disinformation leave a distinctively different digital footprint than those that are found in reality, offering an opportunity to confront such actors through a verification and fact-centred approach to information utilising open-source, social media and digital forensic research that harnesses the power of the digital age. In doing so, the aggressor's actions can be limited by exposing its falsehoods and lifting the veil that covers its crimes and atrocities.

Those seeking to spread disinformation leave a distinctively different digital footprint than those that are found in reality.

Lifting the veil

Open-source footage shows the repeated use of banned cluster munitions and strikes on targets, including mosques, hospitals and water treatment plants in Syria. By comparing and using the masses of information available about these attacks and atrocities, it is possible to examine their number and scale across Syria, the anatomy of individual incidents, and the impact of multiple attacks on individual facilities. This becomes a particularly powerful tool in response to Russia's false claims, lifting the veil of disinformation.

In the final weeks of the siege of the strategic city of Aleppo, Kremlin spokesperson Dmitry Peskov argued that there was no evidence of hospital strikes, and Assad claimed that there was no such policy of targeting. However, the verified proof (including witness testimonies, news footage, videos shot from security cameras and by rescuers, as well as photographs) suggests that the Assad government and its allies, including Russia, did indeed have a policy of targeting Syria's hospitals. For example, the SAMS-supported M2 hospital in al-Maadi district was reportedly damaged in at least twelve attacks between June and December 2016. By examining digital breadcrumbs from the incident (such as open-source videos and images, satellite images of the area around the hospital, and published CCTV footage) it is possible to confirm that the M2 hospital was repeatedly struck between June and December 2016, the damage being consistent with the use of air-dropped bombs and artillery. Equipment and vehicles used by the hospital were damaged and destroyed, and the attacks severely reduced the hospital's ability to serve the local population.

As public awareness of the plight of Aleppo's hospitals grew, so did official denials. Between 28 September and 3 October 2016, the SAMS-supported al-Sakhour hospital (also known as the M10 hospital), was hit in three separate incidents, damaging the hospital buildings and killing staff and patients. In a press conference, the Russian Ministry of Defence (MoD) denied that attacks on the facility had taken place. The MoD briefer, Lieutenant-General Sergei Rudskoy, presented satellite imagery, which he claimed was taken between 24 September and 11 October, stating "no changes to the facility can be observed"

and that "this fact proves that all accusations of indiscriminate strikes voiced by some alleged eyewitnesses turn out to be mere fakes". However, open-source and satellite imagery illustrated different levels of damage to the hospital area after each attack, proving that the Russian MoD's imagery was deceptive[26].

As with hospital strikes, reports of incendiary strikes have been vigorously denied. In late 2015, Major-General Igor Konashenkov, the spokesperson of the Russian MoD, explicitly denied the use of incendiary weapons and accused Amnesty International of "fakes" and "clichés" in a report alleging their use[27]. However, RT (formerly Russia Today) broadcast a striking piece of evidence on 18 June 2016, from Hmeimim, a primarily Russian air base southeast of the city of Latakia. Footage of the Russian defence minister visiting the base showed RBK-500 ZAB-2,5S/M incendiary cluster weapons being mounted on a Russian Su-34, a fighter ground attack aircraft operated only by Russia in Syria[28]. The specific part of the video showing the incendiary cluster weapons was later cut out of a version of the video report uploaded to YouTube by RT[29].

As with the hospital strikes, some of the reported incendiary attacks have been documented in detail and can be independently verified. One such attack occurred between the towns of Rastan and Talbiseh in Homs province on the night from 1 October to 2 October 2016. Local pro-opposition media uploaded a video to their Facebook page that purportedly showed the moment of impact of the incendiary weapon[30]. In the days following the incident, the Syrian Civil Defence—the White Helmets—published photos on their Facebook page claiming to show weapon fragments[31]. Using reference photos and inscriptions on those remnants, the Conflict Intelligence Team (CIT), a group of Russian digital forensic researchers, positively identified the weapon as a RBK-500 ZAB-2,5S/M incendiary cluster bomb[32].

The Cyrillic inscriptions on the casing read RBK 500 ZAB-2,5S/M. ZAB is an abbreviation of the Russian Зажигательная Авиационная Бомба ('incendiary aviation bomb').

Further, weapon remnants resembled reference photos of the cluster and submunitions available from open sources. A large remnant strongly resembled the lid (nose part) and cylindrical casing of an RBK-500 series cluster bomb, and the smaller remnants were identified as two different types of incendiary submunitions: the ZAB-2,5S and the ZAB-2,5(M). These specific types of weapons were not documented prior to Russia's intervention in Syria, leading CIT to conclude that the airstrike was likely conducted by the Russian Air Force. CIT was not able to establish whether the buildings targeted had been inhabited: if they had, the group argued, the attack would have been illegal under the convention[33].

The opportunity

Even though the conflict in Syria rages on and Vladimir Putin managed to keep the international community in a stalemate over how to address the crisis, Russia's disinformation campaign in Syria has also shown weaknesses that serve as opportunities to hold regimes and autocratic governments accountable.

In a hyperconnected age, fighting disinformation by countering disinformation only one event at a time is an approach that brings limited gains and leaves the wider challenge unsolved. Simply countering disinformation by presenting opposing narratives is a symptoms-focused approach, and fails to address the source and methodology of information campaigns. Further, a lack of digital resilience and the lack of government guidance and education to equip policy-makers and citizens with appropriate tools have left societies vulnerable to less benevolent forces that know how to take advantage of such a vacuum.

> *Fighting disinformation by countering disinformation only one event at a time is an approach that brings limited gains.*

What is required is an approach that empowers individuals not only to discover information about Putin's war in Syria, but also to verify the information themselves. Such an approach is the polar opposite

of Russia's opaque disinformation campaign, which relies on ideological narratives over verifiable facts. Western societies must be armed with methods that assist them to differentiate between what is fact and what is fiction.

Only with a robust civil society in place can a credible response unveil the crimes committed by regimes. Adopting hyperconnected solutions around a methods-centred approach to defeating disinformation by actors such as Russia in the Middle East will become more important as the Internet expands. More importantly, as the use of artificial intelligence and deep learning to create disinformation grows, undermining disinformation through a robust level of digital resilience will become increasingly important.

CHAPTER 7

CHINA'S APPROACH TO INFORMATION AND INFLUENCE

Under Xi Jinping China has intensified its efforts to control cyberspace in order to reinforce the domestic rule of the Communist Party and to spread Chinese soft power abroad. Propaganda efforts have been successful domestically in shaping the views of the population, which is isolated from the global Internet. Abroad, China has effectively portrayed itself as a rising power. However, propaganda to promote Chinese foreign-policy objectives on a global scale have not always achieved their objectives.

China has moved into a new phase in its international relations that reflects a growing sense of power and accomplishment, often expressed in terms of reaching the summit or returning to the centre of the world stage. This is expressed by a greater willingness to reject Western norms (or replace them with norms with 'Chinese characteristics') and to assert a larger role for China globally. Domestically, this means tighter and more extensive controls over information. Internationally, it means an effort to garner soft power for China.

China's long standing defensive effort to avoid political risk through information and information technologies, a central inheritance from the Chinese Communist Party's (CCP) Leninist heritage, is now complemented by an effort to reshape global opinion and rules to better serve China's interests and the Party's world-view. The goals of China's information policy are to reduce risks to political stability and continued Party rule; promote Chinese content and technology; reshape global rules to favour China's interests; and defend against perceived US hegemony. Beijing, in the last few years, has created policies and regulations to make the information environment in the country more controllable, most recently with the National Cyberspace Security Strategy released in 2016. China has also become

much more confident in its rejection of universal values, claiming that these are instead 'Western'.

China's leaders see the Internet as an existential threat to stability and continued CCP rule. This view has intensified under Xi Jinping. Xi inherited in 2012 a slow-moving crisis that threatened continuity, and the Xi government has moved forcefully in response. His efforts to ensure economic stability, reduce corruption, reform the People's Liberation Army (PLA), and impose expansive controls on the Internet reinforce his authority and reduce the risk of political instability.

The threat posed by the Internet is also now seen as an opportunity. Since the Chinese Communist Revolution, China has used propaganda and information to control its population, but since Xi has taken office, it now also aims to reach a global audience with this same approach. This reflects the belief that China is on a steady path to becoming the most powerful nation in the world, displacing the US and, therefore, able to extend and perhaps impose Chinese values. Beijing began its pursuit of soft power a decade ago, when former CCP leader Hu Jintao called for making "socialist ideology more attractive and cohesive". Party officials talk about the imminent return of China to the summit of global soft power as it becomes a "powerhouse of discourse" to match its economic power.[34]

Part of China's approach to the threat of information has been to isolate their national networks as much as possible, to build national industries to produce indigenous technologies, and to populate the media with government controlled news and information. China uses censorship and trolls (the '50-cent party') to shape social media in ways favourable to the regimes and damaging to the US. This approach is very effective for domestic audiences, but largely ineffective for foreign ones.

China has a coherent view of cyberspace that places sovereign control by governments at the centre of information policy. It promotes a very different vison of international order that reasserts the primacy of national sovereignty and devalues international agreements that constrain sovereignty, particularly the Universal Declaration of

Human Rights. The country is not alone in this and receives significant support from some non-aligned nations and, of course, Russia. There is a correlation between a nation's willingness to restrict freedom of speech and the likelihood that it is sympathetic to China's views on the Internet and cyberspace.

The emphasis on sovereignty has been accompanied by a major reorganisation of the government and Party apparatus for dealing with cyberspace, including the creation in 2014 of a Central Leading Group for Internet Security and Informatisation, chaired by President Xi, and a new agency, the Cyberspace Administration of China (CAC). Other actions to reinforce domestic control include restrictions on Virtual Private Networks (VPNs) and disruptions to the service they offer, and new limits on social media by deleting posts and closing accounts. The Leading Group sets policy which the CAC implements, improving China's control over domestic networks and Internet users. These changes are the result of a deep interest by President Xi in extending control over cyberspace, which he has identified (along with corruption) as a considerable threat to political stability and CCP rule.

There is a correlation between a nation's willingness to restrict freedom of speech and the likelihood that it is sympathetic to China's views on the Internet and cyberspace.

China uses its World Internet Conference (WIC) to gain support for its ideas of 'cyber sovereignty' and a multilateral approach to Internet governance, but since 2014 (the first WIC) the focus has become more domestic than international. This first reflected the failure of the WIC to attract an influential foreign audience, and reflected greater Chinese confidence in their ability to manage the Internet and extend sovereign control over networks even without being able to expand their control of Internet governance. In general, many Chinese policy-makers believe that the trend in international events favours China, so that they will, over time, achieve their objectives. This may

explain, in part, why the WIC held from 3 to 5 December 2017 drew high-profile technology leaders from around the world.

The Party, not the individual, has primacy. The National Cyberspace Security Strategy asserts that "National sovereignty extends to cyberspace, and cyberspace sovereignty has become an important part of national sovereignty". Xi defined the elements of cyber sovereignty at the 2016 WIC as "respecting each country's right to choose its own Internet development path, its own Internet management model, its own public policies on the Internet, and to participate on an equal basis in the governance of international cyberspace—avoiding hegemony and interference in the internal affairs of other countries".[35] *China's views on sovereignty seeks to reassert the dominant role of states in an approach to globalisation that seeks to amend rules, institutions and standards in ways favourable to its own interests and more consistent with its own political views*.

Beijing has been successful in extending sovereign control to the Internet. It blocks access to and traffic from foreign sites of which it does not approve. Equally important, it shapes the domestic news in ways favourable to the party, emphasising strength, economic growth, China's growing prestige and, recently, the wisdom of Xi Jinping. It is easy to discount the effectiveness of these efforts, and there is a substantial population of Chinese 'netizen's' who mock or express skepticism about the official positions. China uses the full spectrum of media—print, television, film and Internet—to advances its narrative. Survey data from the Pew Foundation and the Chinese Academy of Social Sciences shows that the Chinese public's interest in online content focuses on entertainment, sports and Chinese-source news and that, in fact, the propaganda is effective.

However, the CCP also fears that it could lose control of nationalist sentiment; it is an imprecise tool that Beijing uses with caution. Chinese interlocutors say that social media and 'Colour Revolutions' are a threat, as they could lead to domestic unrest, but believe that the Party is in the process of learning how to deal with and use them for its own purposes, such as by using government employees (the Chinese equivalent of Russian media trolls) to plant millions of

positive messages about the Party and Chinese policies on social media sites[36]. China has found ways to use the IT revolution to extend social control through ubiquitous surveillance in urban areas and online activities.

This sovereign manner is reflected in China's approach to multilateral cybersecurity negotiations, information technology standards, and Internet governance. Its goals are to promote sovereign control and to advance its security and commercial interests. China's new National Cyberspace Security Strategy emphasises "increasingly fierce competition" to "seize the right to develop rules".

The Chinese are cautious and inflexible in international negotiations on cybersecurity in the UN and elsewhere, concerned with defensive requirements, to protect themselves from what they see as a hostile and technologically superior US whose actions are largely untrammeled by international law and are motivated by plans to disrupt Chinese society. China pursues international agreements that would reduce political risk and move in the direction of increasing governmental authority over the Internet. Part of the rationale for opposing norms is a rejection of 'Western' values, but China also blocks agreement on norms that could potentially be used to justify retaliation against China for its *cyber activities*.

Promoting indigenous information technology

Beijing has sought to build a strong information industry since the opening to the West more than three decades ago. This is an important part of its strategy for dealing with cyber and informational risk. China's motives in expanding its IT industry are both commercial and political. China employs various strategies to displace Western IT companies, using non-tariff barriers, security regulations, procurement mandates, and the acquisition (both licit and illicit) of foreign technology, as well as through strategic investments and the acquisition of Western firms.

China has increased its involvement in international standards-setting for information technologies (previously the domain of Western

companies), both to garner commercial advantage and to revise standards, protocols and architectures to improve governmental ability to control cyberspace. Some are calling the race to develop '5G' mobile Internet standards "China's chance to lead global innovation[37]".

> *A senior Chinese official once remarked that if China had not blocked Google from the China market, there would be no Baidu.*

China hopes to repeat the success of Huawei, and use government investments and barriers to entry to produce globally dominant national champions. It has a well-financed strategy to create a domestic industry intended to displace foreign suppliers. A senior Chinese official once remarked that if China had not blocked Google from the China market, there would be no Baidu. Creating a counterpart company and blocking Western services (such as Weibo instead of Twitter) was an effective policy for controlling social media use by a domestic audience, but it is not effective overseas.

Projecting soft power

Chinese propaganda is effective in shaping the views of a domestic Chinese audience, but is far less useful in other countries. China's information operations suffer from a lack of subtlety and attractiveness, and are undercut by China's harsh dealings with its neighbours and its domestic repression. Propaganda has been most effective in persuading the world of its inevitable economic ascendency and in exposing US shortcomings, but it has not succeeded in persuading a non-Han audience that China is an attractive alternative.

Chinese discomfort with the dominance of Western media (such as the BBC or CNN) and their ability to create a global narrative has led China to create competitors to challenge 'information hegemony'. *Global Times* was remade in 2009 to provide English-language content promoting a more positive view of China, complete with its sometimes-shrill, anti-American commentary. Similar views can be found in CCTV (China Central Television), which offers foreign-

language broadcasts in eight major languages, with the explicit goal of creating a more positive narrative of events in China. State-supported Chinese firms have purchased media outlets (such as the *South China Morning Post*) and may reshape reporting and editorial policies along these lines. Executives at Alibaba, the Chinese purchaser, said their goal was to "improve China's image and offer an alternative to what it calls the biased lens of Western news outlets[38]".

Chinese outlets use Western media formats to shape foreign and domestic views of both China and the US in ways favourable to Beijing, even releasing a music video with Chinese rap music interspaced with official pronouncements to extoll Xi and the 19th CCP Congress—even the opening words are in English. While these information operations are very effective in influencing the views of a Chinese audience, they are much less successful in other cultural and linguistic arenas. A gaming app that allowed users to use a smart phone to 'clap' for President Xi went viral in China but received little notice overseas.

China has taken both a hard and soft approach to engendering a degree of self-censorship among Western firms, which do not wish to alienate Beijing or lose market access. Western film producers are careful not to offend Chinese censors (such as when the army invading the US in the remake of Red Dawn was suddenly changed from the PLA to North Korea's, or when China saves NASA in The Martian). Shows that portray the US in a negative light, such as Netflix's House of Cards are permitted for rebroadcast in China (and many Chinese saw it as a quasi-documentary).

How effective these efforts have been in reshaping foreign views of China is open to question. It is too early to assess the effect of the country's media purchases, but when Alibaba purchased the *South China Morning Post* it was with the explicit goal of creating more positive coverage of China. The creation of Confucius Institutes, a heavy-handed effort at soft power in the US, where most of the Institutes are located, had mixed results, attracting criticism from a range of sources without noticeable improvement in US views of China[39]. Similarly, Chinese efforts to influence Australian views,

using political donations and student or immigrant organisations. China's message remains most attractive to Chinese nationals resident in other countries.

The Chinese do not have doctrine to create 'cognitive effect' and disinformation similar to what has been developed by Russia. China appears to rely on extending techniques developed for domestic control to foreign audiences. An initial assessment is that Chinese efforts have been more effective on the country's own population. Beijing has not been able to devise an attractive alternative. Its own ideological constraints, which increasingly contain elements of the personality cult seen under Mao, are unpersuasive to non-Chinese audiences. A mixture of domestic coercion and financial pressure on overseas audiences remains China's most effective tools for influence.

A mixture of domestic coercion and financial pressure on overseas audiences remains China's most effective tools for influence.

In looking at all these activities, they point to a coherent strategy to control information, centrally developed and overseen, to minimise political risk, and advance a Chinese agenda and narrative internationally. The Chinese state sees information and information technology as a tool in ways not found in Western democracies.

CHAPTER 8

FROM 'LIKES' TO LEADERS: THE IMPACT OF SOCIAL NETWORKS IN THE PHILIPPINES

Social news network Rappler.com has documented the latest presidential campaign in the Philippines. A highly-targeted social media campaign was instrumentalised to support the election of Rodrigo Duterte, then was turned against the president's critics, opposition leaders and the traditional media. The government has thus succeeded in suppressing independent voices in favour of government messages.

Patriotic trolling, which an international research coalition[40] defines as "the use of targeted, State-sponsored online hate and harassment campaigns leveraged to silence and intimidate individuals" is operating in the Philippines[41]. With nearly 97 per cent of the Filipino population on the Internet using Facebook, the vulnerability of the Philippines to such campaigns has been identified and openly exploited.

Rappler journalists and data scientists have documented hundreds of web sites and millions of social media accounts and groups that methodically and consistently spread disinformation in the Philippines—culminating in a database of more than 11 million personal profiles and 250 million public comments (as of March 2017). This work has uncovered the emergence and evolution of a complex patriotic trolling network aimed at electing and supporting Rodrigo Duterte, the winner of the 2016 presidential elections.

To get a sense of the network's reach and power, Rappler spent three months manually tracing a sample 'sock puppet network' of 26 fake Facebook accounts. These accounts were found to have influenced up to three million Facebook users. In addition, in November 2016, Rappler documented more than 50,000 accounts on Facebook that were under the direct control of the propaganda network, including fake accounts (some clearly centrally managed), paid trolls, and real supporters working to convince their families and friends. By April

2017, clear links with the state began to appear, most notably the office in charge of state media under Secretary Martin Andanar, the Presidential Communications Operations Office (PCOO).

By mid-2017, patriotic trolling formed the foundation of the Philippine government's information ecosystem, discrediting institutions, politicians and journalists who questioned or criticised its actions. This ecosystem's priority is to defend President Duterte, now the most powerful Filipino leader in the last three decades, and his high popularity ratings. He controls a supermajority in the legislature, will appoint 13 of 15 Supreme Court justices, and has essentially dismantled any effective opposition.

Evolution of the machine and its targets

The first social media campaign to successfully elect a president in the Philippines tapped into collective and justifiable anger between economic classes. This campaign network was instrumental in electing the nation's leader, Rodrigo Duterte. Broken into four different geographical groups, the distribution network on Facebook received daily messages from a central messaging group that worked with psychologists to design messages that would appeal emotionally for viral spread. Ironically, the social media networks created during the campaign were weaponised only after Duterte was inaugurated on 30 June 2016. The President then decided to boycott traditional media for approximately one month, triggering the second phase. In this phase, the network evolved, using more targeted and virulent strategies, which transformed existing campaign-based social media accounts to accounts meant to attack opposition leaders and traditional media. Harnessing its massive base, it acted to successfully stifle dissent and shape public opinion about controversial policies like President Duterte's drug war, conspiracy theories, foreign policy, martial law and other government initiatives.

President Duterte's goal was clear and effective: tear down the credibility of anyone questioning or critical of the government. By making an example of one citizen, one politician, one journalist, all brutally attacked online, it created a chilling effect that made many

others afraid to speak out. One of the first targets was Senator Leila de Lima, former justice secretary and former head of the Philippines Commission on Human Rights. The attack on the senator was followed, in January 2017, by the targeting of other female politicians, including Vice President Leni Robredo and Senator Risa Hontiveros.

> *The first social media campaign to successfully elect a president in the Philippines tapped into collective and justifiable anger between economic classes.*

ABS-CBN, the country's largest television network, and the *Philippine Daily Inquirer*, the largest newspaper, were the first media targets in an effective campaign that pushed to tone down critical reporting. The *Inquirer* was targeted for its 'Kill List', its roster of people killed during the drug war. Shortly after the concerted attacks, the *Inquirer* abandoned maintaining the list, and both news groups backtracked on the number of people killed. Rappler maintains that, based on figures released by the police, about 7,000 people were killed in the drug war from 1 July 2016 to 31 January 2017. This amounts to approximately 1,000 people killed per month. After growing international condemnation, the Philippine government began to blur the actual numbers, changing its definitions and including deaths under investigation (DUI) as a new category created by the police.

This was followed by one of the most publicised outreach programs by the Presidential Palace or PCOO. Dubbed #RealNumbersPH, the government actively worked with bloggers from the social media propaganda machine to pressure traditional media to change their numbers to the new "official numbers". During these months, any time anyone on Facebook brought up the rising death toll in the drug war, that person would be viciously attacked. The end goal was to silence criticism, effectively creating what mass communications theory calls a "spiral of silence[42]".

Journalists and news groups, which once held the highest credibility ratings among public and private institutions in the Philippines, were

systematically attacked and degraded, first on social media, then by government officials (including President Duterte). Many of the same themes that first appeared in the election campaigns were carried over and amplified: that journalists are corrupt; that news organisations are owned by oligarchs with vested interests; that clickbait headlines brought their own economic gains, etc. In 2016, President Duterte publicly and repeatedly threatened ABS-CBN and the *Philippine Daily Inquirer*.

> *By making an example of one citizen, one politician, one journalist, all brutally attacked online, it created a chilling effect that made many others afraid to speak out.*

Patriotic trolling first focused on Rappler and its CEO after the company published a three-part social media propaganda series in early October 2016. Backed by data, it was the first time the full scope of the propaganda machine was unveiled publicly. The machine immediately retaliated, calling for attacks against the Rappler CEO that reached as many as 90 hate messages per hour[43] and a #UnfollowRappler campaign on social media that exposed the extent of its power in the virtual world.

By November 2016, online behaviour and data showed that the machine could command and influence a little more than 52,000 accounts, a significant number when compared to the 30,000 accounts that Facebook shut down in the lead up to the French elections[44]. Incidentally, Facebook later noted that its work during the French elections was shaped partly by the data Rappler had provided them as early as August[45].

Breaking down trust

The third wave of attacks began in early January 2017, first targeting Vice President Leni Robredo and other women leaders using half-truths, outright lies, sexism and misogyny.

Women are favoured and effective targets attacked, derided and ridiculed, often with demeaning sexual slurs and curses. This near-constant onslaught further polarised Filipino society and deepened the spiral of silence.

Social media accounts supporting and allegedly funded by the government actively worked to cripple trust in what was then a virtually non-existent opposition, and in journalism and other credible sources of information, working to replace them with the government's voice amplified through social media. Fake news sites grew from 15 to more than 300 in a few months, spread by fake accounts, bots and 'keyboard warriors' sowing confusion and distrust, and leaving government with the loudest megaphone.

By February 2017, the propaganda machine focused on Rappler in near-daily attacks attempting to paint the start-up as foreign-owned or controlled by foreign interests in order to influence events in the Philippines. Despite repeated denials, many Duterte supporters believed the narrative repeatedly pushed by pro-Duterte bloggers, and a claim would be repeated several months later by President Duterte himself in his annual State of the Nation Address.

By mid-year, attacks on the media intensified. President Duterte again publicly attacked ABS-CBN and the *Philippine Daily Inquirer* while the propaganda machine attempted to trend #ArrestMariaRessa and paint Rappler as a tool for foreign intervention on social media.

State-sponsored attacks

By this time it was clear that the online propaganda machine was the harbinger and test site for government messages and attacks against its perceived critics. Rappler identified three key content creators of the propaganda machine, which segmented Filipino society by economic demographics: Sass Sassot for the pseudo-intellectual posts for the top one per cent; Thinking Pinoy (RJ Nieto) for the middle class; and Mocha Uson for the mass base.

The government closed the loop by bestowing Mocha Uson and RJ Nieto with government positions. Mocha Uson was appointed assistant secretary in charge of social media under PCOO; RJ Nieto is employed with the Department of Foreign Affairs and the Department of Transportation (DOTR). Their networks are also the government's first line of alert and defence in crisis management. On 23 May 2017, the Philippine government declared martial law in Mindanao, changing the landscape significantly. The announcement was made from Moscow during a state visit to Russia, which included both Mocha Uson and RJ Nieto, and helped set the stage for the fourth wave of attacks, combining online and real world government actions to limit press freedom.

On 17 July, the *Philippine Daily Inquirer* called a general assembly and informed its staff that it would be selling the newspaper to Ramon Ang, a businessman with close ties to President Duterte. This development came after cases were filed against the family that owned the Inquirer, board members were threatened with tax cases, and an informal advertising boycott plunged its revenues by at least 40 per cent.

One week later, at the annual State of the Nation Address, President Duterte attacked Rappler[46], along with ABS-CBN and the UN, Barack Obama, the ICC and others. He would repeat the attacks against Rappler on three more occasions in the following three weeks. Incidents of harassment began that same week, with one of the pro-Duterte bloggers releasing all of Rappler's financial statements on Facebook. This was followed by unprecedented requests and calls from the Security and Exchange Commission, which began a special panel investigation.

Role of US technology giants and the road ahead

The irony, of course, is that the greatest threat to democracy in the Philippines is enabled by US companies: Facebook[47], Google and Twitter. YouTube, the world's second largest search engine operated by Google, is also a favourite and an effective platform for video attacks. The explosion of information and the black box of algorithms

has demolished journalism's role as gatekeeper, shifting the collective narrative from human editors to machines and algorithms.

The latest reports and analyses show that this rollback of democracy is occurring in at least 30 countries around the world, according to a November report by Freedom House[48]. In the short-term, the solution to protect democracy is in the hands of these US companies as they learn to deal with the impact of the complex systems they have created. The medium-term solution is in greater media literacy and an acknowledgement of this world of exponential information lumping together truth and lies. In the long-term, it is education.

> *The medium-term solution is in greater media literacy and an acknowledgement of this world of exponential information lumping together truth and lies.*

Tech giants need to build democracy into their algorithms and prevent autocratic governments from successfully building online armies. A difficult proposition when the platforms' competing economic interests and mandates for growth are considered.

CHAPTER 9

COUNTERING DISINFORMATION IN UKRAINE

StopFake.org began as a vehicle to refute Russian fake news stories about Ukraine, and has now turned into an international information hub on Kremlin propaganda. Its team of journalist has launched numerous tools for debunking Russian narratives, discrediting Russian propaganda and conducting education programs to increase media literacy.

StopFake.org is a fact-checking project that tackles Russian disinformation and propaganda by debunking fake news. Launched in 2014 by journalism professors, students and alumni of the Mohyla School of Journalism in Kyiv, it was a reaction to the annexation of Crimea and Russia's war against Ukraine in the Donbass region. Initially, the goal of the project was to verify and refute disinformation and propaganda about events in Ukraine being circulated in the media. The project has grown into an information hub where all aspects of Kremlin propaganda are carefully examined and analysed.

To date, the organisation's team of 30 has debunked more than a thousand stories from Russian mainstream media (TV channels, newspapers, news agencies) in 11 different languages. The content—which includes text video, audio content, syndicated television and radio shows, a local Donbass newspaper, and a documentary—reaches 230,000 followers on social media and numerous others in person. As the holder of the largest archive of Russian fake news, StopFake.org fact-checks, de-bunks, edits, translates, researches and disseminates information.

Monitoring, debunking, archiving and defining main narratives

Russia's war against Ukraine illustrated the contemporary use by Moscow of propaganda globally, but as a tool, this represents a

continuation of Soviet methods—adapted to increase its impact and efficiency today.

Television remains as one of the main instruments of influence and dissemination of disinformation in Russia and abroad. The value of this medium to Russia is expressed by Margarita Simonyan, RT head: "To some extent, if you do not have broadcasting for abroad - it's like you do not have the army. When there is no war—you do not need it. But when the war [has] already started, you cannot create it in a week[49]." Well before the beginning of the Crimean annexation, Russian television was a significant channel for influencing Ukrainian public opinion, with all major channels freely available in Ukraine and Ukrainian state-run technical facilities being used for carrying and amplifying signals. Russian TV content was widely consumed in Ukraine as a result of the linguistic proximity and a partially integrated media economy between the two countries.

> *As a tool, this represents a continuation of Soviet methods—adapted to increase its impact and efficiency today.*

At the same time, other segments of the Russian media system were also dominating over the Ukrainian media landscape, including Internet news media, social media and the shared entertainment industry. All were gradually weaponised, with the Russian media involved in manufacturing and distributing textual fakes, manipulative titles, visual fakes, false claims, forged documents, phoney experts, fake news sources and witnesses. Together, they culminated in a series of fake narratives discrediting different aspects of life in Ukraine, which were then targeted at audiences in Russia, Ukraine and globally. To maximise the disinformation effect, narratives were repeated, translated and amplified by social media.

StopFake.org's primary objectives include debunking fake narratives, disseminating findings to different audiences and building an archive of cases. Preliminary analyses of 500 items of disinformation produced by Russian propaganda on Ukraine during the period 2014-2015

identified multiple major manufactured narratives, including the following:

- Depictions of Ukraine as a fascist and failed state whose territory was in constant disintegration, dispute or threat of annexation by neighbouring and Western countries;
- Manipulations of political or economic relations with international partners, including de-legitimisation of the EU and NATO and misrepresentations of foreign support of, and intentions in, Ukraine; and
- Characterizations of Russia as not participating in the affairs of Ukraine, including denials of a Russian occupation and involvement in the crash of flight MH17.

Further analysis regarding Ukraine revealed that the largest number of fake news stories (79 items) was generated by Zvezda TV, which belongs to the Ministry of Defence in Russia. The second largest producer with 73 items, Ukraine.ru, is a Russian web site belonging to the Russian state-owned Novosti information agency, followed by the RIA Novosti information agency with 62 fake items.

> *The entire Russian media landscape serves the Kremlin's objective of manufacturing and distributing fake news.*

Both state-owned and privately-owned (but state-controlled) media are involved in Russia, with television and the Internet dominating the propaganda ecosystem. Importantly, analyses illustrate that the entire Russian media landscape serves the Kremlin's objective of manufacturing and distributing fake news. This system is a major component of Russia's information warfare in Ukraine.

Discrediting Russian *agitprop* and raising domestic and international awareness

Russian propaganda operates beyond the Russian language realm and is active on a global scale. While Russian speakers are more likely

to follow Russian domestic media, RT (formerly Russia Today) operates in five languages and Sputnik in 31; non-attributable web sites and trolls operate in many other countries and in a multitude of languages. The reach of Russian propaganda is therefore not constrained by language or location, making public awareness a top priority.

Research in early 2017 by StopFake.org illuminated the perception of Russian propaganda by Ukrainians and their resilience towards it:

- The majority of Ukrainian citizens (58.3 per cent) share the opinion that there is a threat of Russian propaganda in Ukraine;
- Ukrainians view Russian TV channels, online media and social networks as the most widespread sources of Russian propaganda (45 per cent, 34.5 per cent and 19.8 per cent respectively);
- The majority of the Ukrainian population (59.7 per cent) believes that they are able to distinguish truthful information from false information in the media; and
- 42.1 per cent of respondents believe that disinformation is a serious problem in Russian media.

An important step in disconnecting Ukrainians from Russia's propaganda pipeline was the removal from air of 75 Russian TV channels previously available in Ukraine. Decreed by a Ukrainian court in 2014 at the beginning of the war in Donbass, the removal resulted in a dramatic drop in Russian TV news viewership in Ukraine, from 12 per cent in 2015, to 7 per cent in 2016, and 5 per cent in 2017.

The shrinking Russian media audience in Ukraine can also be explained by limits imposed on the presence of Russian social media companies. In May 2017, the President of Ukraine (Poroshenko) signed a decree blocking Russian social networks from operating in Ukraine as part of a wider set of sanctions. The inability for Ukrainian Internet service providers to provide access to Russian social networks had a tremendous impact: according to SimilarWeb, the Ukrainian audience of VKontakte decreased by 60 per cent in 2017, from 9.8

million to 3.8 million visits per day, while visits to Odnoklassniki ('Classmates') fell by 64 per cent (from 4.6 million to 1.6 million visits per day). Both social networks were known to host thousands of anti-Ukrainian groups and disseminate propaganda, and were operational tools to raise funds and hire mercenaries for the war in Donbass.

The drop in the number of visitors to the Russian search engine Yandex, which provides a spectrum of personalised services and harvests geolocation and other data from Ukrainian users, reached 65 per cent, from 5.9 million to 2 million visits per day. Mail.ru—one of the most widely used email service in Ukraine—lost 55 per cent of its Ukrainian audience. Many of those users were Ukrainian military members who were regularly targeted with Russian manufactured news through the ads section of this service.

Disseminating knowledge and promoting media literacy

In Ukraine, StopFake.org also works to improve media literacy of different audiences, with a special focus on the populations of Donbass and Crimea (despite the obvious difficulties in reaching these audiences).

In 2015, StopFake.org conducted media literacy training for general audiences in eastern and southern Ukraine. The project consisted of 'training the trainers', curriculum and training manual development, and a series of intensive one-day training sessions for targeted audiences determined to be at risk from Russian propaganda. The training was accompanied by an intensive advertising campaign in the national and local media (TV, radio, banners on news web sites and social media, and outdoor advertisements) promoting media literacy and providing tools for citizens could use to check facts. As a result of this project, more than 15,000 individuals were trained in the basic skills needed for more critical media consumption.

Ukrainians continue to face difficulties grasping the challenges of a post-truth era. According to polling conducted in February 2017, most participants, especially of the middle and younger age groups, have

heard of and understand the concept of a 'fake' when applied to news. Nevertheless, the concept remains unusual for many. All participants, even the youngest cohort, noted that they do not use it in everyday parlance and consider it slang used by young people and teenagers. In contrast, the concept of propaganda was clear for most participants, especially those in the middle and older age cohorts who were politically aware during Soviet times. Considering that younger audiences are more likely to use social media platforms, these findings highlight a critical need for further media literacy training.

To expand its work internationally, StopFake.org partners with many fact-checking organisations and networks across Europe to share the Ukrainian experience, raise global awareness of Russian disinformation and its influence on political processes and decision-making, as well as facilitate political discussions of disinformation in other countries.

CHAPTER 10

FAKE FOR PROFIT: NON-STATE ACTORS AND THE BUSINESS OF DISINFORMATION

Fake news entrepreneurs profit from click-based advertising directed at readers of sensationalist stories and those who limit their news consumption to online news aggregating web sites. These enterprises maximise their readership and clickbait potential by purchasing the pages of groups with sizeable memberships which fit the target demographic. The truth, falsehood, and subject matter of their news content are irrelevant—the singular objective is attracting readers who will view advertisements.

This paper is primarily based on a lengthy one-to-one interview with a Kosovan called Burim. Twenty-four years old, Burim (not his real name) had graduated with a degree in computer science. He had worked in IT for a private company in Kosovo's capital, Prishtina, and in advertising. Since January 2016, Burim has been the owner-operator of an online spam and disinformation operation.

Conducted in Kosovo in June 2017, the interview was part of a wider attempt to understand the phenomenon of disinformation through gaining an appreciation for the life, motivations, beliefs and anxieties of someone like Burim. The production of disinformation is a phenomenon that, doubtless, is intimately related to the technologies that allow for the publication and consumption of content. But it is also something that humans decide to do, and it is hoped that this contribution helps inform an understanding of why they do so.

The audience

The preliminary objective of Burim's operation is to capture attention, and the sole platform he employs for this purpose is Facebook. At any time, he 'owns' approximately one dozen Facebook pages. One appears to be an evangelical group, with a big picture of Jesus Christ. "I bought this one" he noted. "This guy in Albania built up this page

by posting authentic religious information. He managed to get 100,000 likes on the page. Then I paid him 2,000 euros, and he transferred the page over to me." Another page is about abandoned places, and another about mobilising communities in a city in the south of the US. One he had bought just recently, originally a group dedicated to sharing tips and information about dieting and veganism. There was a group about tiny houses and another was a verified page—it had a blue tick, and a logo—that had something to do with trust. It was quite difficult to actually see what most of Burim's pages had originally been about. But while the groups were bizarre, their audiences were huge: 90,000 likes, 240,000 likes, 26,000 likes. In Burim's quest to develop an audience, these pages could, at least in theory, present his content to close to one million individuals.

He acquired the groups in different ways. He had a centrepiece page that he had built himself, investing 20,000 euros into targeted advertising on Facebook to build the audience of the page to just over 100,000 members. It was the most honest of any of the pages he owned, explicitly dedicated to sharing the day's viral, trending stories. But most of the groups, he purchased. In some instances, Burim approaches the administrator of a group directly to explore if they are willing to sell it, "if I come across something interesting, I'll try to buy it". But most of the groups were purchased from an informal network of people who themselves bought and sold pages, predominantly also for the purpose of producing clickbait and spam.

"We don't know if the groups will work beforehand" Burim explained, "so we post some content and wait three or four hours to see how many people are clicking on it. That's how we know whether a page is going to be helpful." Burim and his team test each group that they have newly acquired, checking the scale of clicks and shares that their content generates. The targeted Facebook users are "digitally illiterate, preferably Americans and usually 30 years old or older." Deliberately avoided are groups with audiences that are too young, and any groups that are hypothesized to have too many technologically savvy members. "We need to reach people who don't understand the digital world or clickbait." If the content does not generate traction,

the group is quickly sold onwards in order to free the capital to invest in another group.

The content

Burim employs seven people to keep the content flowing through his groups. Their job, however, is not to write any of the content themselves. There is no economic stake is creating content when it can be so easily stolen from elsewhere. Instead, they identify and appropriate content that has already been highly shared, usually from the countless other operations that are similar to theirs. Tracing the origin of most of the stories that they share is incredibly difficult. As the content moves from outlet to outlet, it often gets subtly changed, sometimes shortened, sometimes exaggerated or simplified. Burim describes it as a washing machine—both because the content is never at rest, but also because each 'wash' seemed to slightly change the story, sometimes shrinking, distorting, exaggerating or further warping it, until its origin becomes unknowable and irrelevant.

His target audience is not particularly interested in political content. "Stories about killing people, gore, basically, they perform best!" said Burim, cheerfully. Under his thumb, story after story flicked by. "Dog Groomer who Kicked Dog all its Ribs Broke Remains Jail-Free" was one story. ''Boy Comes out of Coma after 12 Years, Whispers Dark Secret to Parents [video]" was another. "Burn Bay Leaves in your Home for these 13 Amazing Health Benefits"; "The Peanut Butter test—the Easiest way to Detect early Alzeihmer's. Everyone must watch this!" Some had been shared only hundreds of times across his groups, but many were in the thousands and a few in the tens of thousands. In Burim's own eyes, he was giving people what they wanted to click on, content that spoke to his audience's hopes, anxieties, guilty pleasures, and temptations: a desire to be healthy (through easy tricks and tips); to be outraged at (clearly signposted) evil. The content shared by his groups was a procession of the ridiculous, the tragic and the gory.

To call this activity 'fake news' misses the true crux of this phenomenon. The stories are not deliberately false; they are just not

deliberately true. The only thing that matters is the size of the audience that the content can harvest. "I don't care what the group does", he said. "I don't even read it. This is the first time I've actually read it. It's all total nonsense." True, false, the content did not matter. "I don't care what the content is", he said, again, still scrolling through the endless content that his operation spews out. He pauses for a moment, his thumb hovering over a story going crazy, its shares spiking skywards, from the tens into the hundreds of thousands. "I just care about traffic."

> *The only thing that matters is the size of the audience that the content can harvest.*

The money

When the audience clicks on any of the stories that this team posts, they are taken to the moneymaking part of Burim's operation. He maintains around a dozen web sites outside of Facebook and changes the URLs to avoid detection. They look like crude versions of an online newspaper, with the full stories hosted under sections called, variously, 'Home', 'Health', 'DIY', 'Animals', 'Food Art' and so on.

The rise of programmatic advertising has opened up a huge opportunity for people like Burim. Programmatic advertising is an alternative to traditional brand advertising through broadcast media channels. It uses software to buy advertising space wherever a member of a target audience appears on the Internet, often identified through cookies, device IDs or by specialist ad-technology providers. The point is not to sell advertising space on a web site, let alone a newspaper, but space in front of a targeted individual, wherever they happen to be. This has meant that Burim did not have to try selling advertising space directly to agencies. He could sell it through programmatic advertising intermediaries, and just like any (legitimate) newspaper, he earned most of his money through Google AdSense, pay-per-click advertising.

Burim's operation is earning anything from 400 to several thousand euros per day; good money anywhere, and a very substantial income

in Kosovo. He brings an entrepreneurial, business mind set to the operation. The language that he uses to talk about his decisions was that of calculated risk, investment and reward. Some of his groups had been closed down, but these were losses he just shrugged off as occupational hazards.

Future trends

The business environment was becoming tougher, he said. There are at least 200 or 300 people engaged in similar enterprises across Kosovo, Macedonia and Albania. Burim saw himself as an early mover into the industry, but with the volume of competitors increasing, he is finding it more difficult to get the clicks with so many others also vying for their attention.

As in so many other areas, there has recently been a profusion of small, agile actors: fake news start-ups. A small number of players are getting bigger and others are dying out. "I expect it to consolidate", he said. He also knows that Facebook is working to throttle the endless stream of clickbait and run him out of business. To him, this is just another occupational hazard.

Both the identification and publication of content is still predominantly a manual process, and if actors do become larger and better resourced, both will likely become more automated and data-driven. Technologies have been developed for legitimate journalistic outlets (such as BBC Trending or Buzzfeed) to identify quickly stories that are being widely shared, or even (through metrics such as 'viral uplift') that are likely to be widely shared in the future. It is easy to see how enterprises like Burim's might use these technologies to seek advantage over their competitors in finding and re-publishing the most shareable, viral content.

Conclusions and counter-measures

Burim is in many ways the nemesis of good journalism. To him, the content is irrelevant, the provenance unimportant, the story recycled, and the truth not even worth thinking about. But he is also only the product of much more general forces that have swept through

mainstream journalism as well as enterprises like his. Of course television is still the main source of news for those over 55, and most use a mix of different sources to understand what is happening in the world.[50] However, the Internet is now the main source of news for more people than any other medium, and for those that do use the Internet to find their news, most access it indirectly. They use a gateway, from search engines and aggregators, to social media sites and voice-controlled digital assistants. These timelines are often algorithmically curated, and these algorithms attempt to serve up calculated and specific content that the reader would be most likely to engage with. Most of all, the rise of programmatic advertising means that clicks are the way that revenue is earned. Thrown side-by-side into a feed, ranked by engagement and clicks, the risk and cost of good journalism is becoming detached from the payoff in actually doing it.

The risk and cost of good journalism is becoming detached from the payoff in actually doing it.

In the West, poor quality online information is seen as something that poisons political debate and undermines good journalism. However, meeting Burim illuminated another side of online disinformation. The interpreter leaned over: "His accent is from Lipjan", she said, "rural, working class". What Burim is doing might be injurious, even dangerous, to public life, but to him, it is also an opportunity for social mobility. It is a way out of rural poverty, the best prospect in a place where there are far too few.

It is understandable that, seen as a technical problem, technical approaches are viewed as the solutions to online disinformation. However, it is also a social and economic problem. Both technology companies and governments should consider ways to harness and re-purpose the enterprise and intelligence of people like Burim into more socially beneficial and pro-social activities.

ENDNOTES

1 An alternative to traditional brand advertising through broadcast media channels, programmatic advertising involves targeting individual consumers via cookies, device IDs, and algorithmic software, automating the sale of advertising using real-time-bidding.

2 Voelz, J. (2017), "Transnationalism and Anti-Globalism", *College Literature*, 44(4), pp. 521-526.

3 Stringer, J. (2017), "Why did anti-globalisation fail and anti-globalism succeed?", Open Democracy; accessible at https://www.opendemocracy.net/jacob-stringer/why-did-anti-globalisation-fail-and-anti-globalism-succeed.

4 Antifa is a loosely organised, far-left, political action group that claims to be 'anti-fascist'. The group has staged several protests and counter-protests (to alt-right political activities) in 2017 and has been criticised for using violent tactics. As with its alt-right counterparts, some of its online activities have been connected to Russian information operations.

5 Zannettou, S. et al, (2017), The Web Centipede: Understanding How Web Communities Influence Each Other Through the Lens of Mainstream and Alternative News Sources. *Available: https://arxiv.org/pdf/1705.06947.pdf.*

6 Starbird, K. (2017), Examining the Alternative Media Ecosystem Through the Production of Alternative Narratives of Mass Shooting Events on Twitter. *In Proceedings of the Eleventh International Conference on Web and Social Media* (pp. 230-239), ISBN: 978-1-57735-788-9.

7 Pomerantsev, P., and Weiss, M. (2014), "The Menace of Unreality: How the Kremlin Weaponizes Information, Culture and Money", *The Interpreter*, Institute of Modern Russia.

8 Giesea, J. (2015). "It's time to embrace memetic warfare", NATO Stratcom COE, *Defence Strategic Communications Journal*, 1 (1).

9 Culminating in a decree from Andropov dated 12 April 1982, ordering all KGB foreign intelligence officers, regardless of their actual assignments, to engage in active measures with the aim of ensuring that Ronald Reagan would lose his campaign for re-election. See: *The Sword and the Shield: the Mitrokhin Archive and the Secret History of the KGB*, Andrew, Christopher M. 1999, p. 242. Basic Books, New York.

10 Blake, William 1803 (approximately), "Auguries of Innocence"; accessible at https://www.poetryfoundation.org/poems/43650/auguries-of-innocence. Credit for this observation: Russian disinformation campaign: What it takes, CNN, October 2017.

11 Galeotti, Mark, "What Exactly are 'Kremlin Ties'?", *Atlantic Monthly*, July 2017.

12 Michman, Gable, & Gross. *Market Segmentation: A Selected and Annotated Bibliography*, American Marketing Association, Chicago IL, 1977.

13 Mihkelsoni, Marko, *Disinformation across ages: Russia's old but effective weapon of influence*, Euromaidan Press, July 2017.

14 Pacepa & Rychlak *Disinformation*, p. 39, WND Books, 2013, citing Great Soviet Encyclopedia, State Scientific Publishing House, 1952.

15 Hill, Fiona and Gaddy, Clifford G., "How the 1980s Explains Vladimir Putin", *Atlantic Monthly*, February 2013.

16 Kross, Eerik-Niiles, "America, welcome to the war", Politico EU, August 2016.

17 Testimony of Robert M. Gates, Deputy Director for Intelligence, CIA, before the US Senate Foreign Relations Committee, Subcommittee on European Affairs, 12 September 1985.

18 The operations described here, while hypothetical in the strictest sense of the word, bear a striking resemblance to actual Kremlin information operations directly observed and investigated by the author.

19 Translated online from https://rusemb.org.uk/press/2029. Paragraph 15a lists the characteristics of modern conflicts.

20 Davies, Katie, "Revealed: Confessions of a Kremlin Troll", *Moscow Times*, 18 April 2017; accessible at https://themoscowtimes.com/articles/revealed-confessions-of-a-kremlin-troll-57754.

21 "Why the 'fake rape' story against German NATO forces fell flat in Lithuania," DW, 23 February 2017; accessible at http://www.dw.com/en/why-the-fake-rape-story-against-german-nato-forces-fell-flat-in-lithuania/a-37694870.

22 Inglehart, R. F., & Norris, P. (2016), "Trump, Brexit, and the Rise of Populism: Economic Have- nots and Cultural Backlash"; paper presented at the American Political Science Association Annual Meeting, Philadelphia, USA; accessible at https://research.hks.harvard.edu/publications/getFile.aspx?Id=1401.

23 For example, #voteleave, #voteremain, #votein, #voteout, #leaveeu, #bremain, #strongerin, #Brexit, #euref.

24 The Brexit referendum was held on 23 June 2016.

25 Aikins, Matthieu, "Whoever Saves a Life", *Matter*, 15 September 2014; accessible at https://medium.com/matter/whoever-saves-a-life-1aaea20b782#.b60t2sth9.

26 Higgins, Eliot, "Fact-Checking Russia's Claim that It Didn't Bomb Another Hospital in Syria", *Bellingcat*, 9 November 2016; accessible at https://www.bellingcat.com/news/mena/2016/11/09/fact-checking-russias-claim-didnt-bomb-another-hospital-syria/.

27 Amnesty International Report on "Civilian Deaths' Based on Fakes, Clichés", *Sputnik News*, 23 December 2015; accessible at https://sputniknews.com/middleeast/201512231032213565-amnesty-intl-report-fake/.

28 RT's YouTube video, 18 June 2016; accessible at https://www.youtube.com/watch?v=dNbIRD8Cq48&feature=youtu.be&t=44; Leviev, Ruslan "Sputnik, RT and Russian MoD Expose Cluster Bombs at Hmeymim Airbase", *Conflict Intelligence Team*, 7 June 2016; accessible at https://citeam.org/sputnik-rt-and-russian-mod-expose-cluster-bombs-at-hmeimim-airbase.

29 Dearden, Lizzie, "Russia-Backed Broadcaster RT Cuts Footage Proving Use of Incendiary 'Cluster Bombs' in Syria", *The Independent*, 21 June 2016; accessible at http://www.independent.co.uk/news/world/middle-east/russia-today-syria-war-cluster-bomb-footage-censorship-video-vladimir-putin-a7093141.html.

30 "Motasem homs"'s YouTube video. 1 October 2016; accessible at https://www.youtube.com/watch?list=PL3vE7Lp4BcaFpsYlpO92RwSoIj83BnFIq&v=MZY7UvrnxUw.

31 Syrian civil defence in Homs's Facebook post, 2 October 2016; accessible at https://www.facebook.com/SCD.HOMS/posts/603882693122910.

32 "New Evidence of Russian Incendiary Bombs Use in Syria", Conflict Intelligence Team; accessible at https://citeam.org/new-evidence-of-russian-incendiary-bomb-use-in-syria/.

33 For a discussion on the legal aspect, see for example this dispatch on incendiary bomb use in Syria: "Syria/Russia: Incendiary Weapons Burn in Aleppo, Idlib," *Human Rights Watch*, 16 August 2016; accessible at https://www.hrw.org/news/2016/08/16/syria/russia-incendiary-weapons-burn-aleppo-idlib.

34 Bandurski, David, "*China, Rhetorical Giant on the Move*", China Media Project, 22 June 2017; accessible at http://chinamediaproject.org/2017/06/24/china-rhetorical-giant-move/.

35 Ministry of Foreign Affairs of the People's Republic of China, 16 December 2015: accessible at http://www.fmprc.gov.cn/mfa_eng/wjdt_665385/zyjh_665391/t1327570.shtml.

36 Farrell, Henry, "The Chinese government fakes nearly 450 million social media comments a year. This is why", *Washington Post*, 19 May 2016; accessible at https://www.washingtonpost.com/news/monkey-cage/wp/2016/05/19/the-chinese-government-fakes-nearly-450-million-social-media-comments-a-year-this-is-why/?utm_term=.9d718382c7fd.

37 "China's chance to lead global innovation may lie with 5G mobile technology development", *South China Morning Post*, 1 October 2017; accessible at http://www.scmp.com/tech/enterprises/article/2113581/chinas-chance-lead-global-innovation-may-lie-5g-mobile-technology.

38 Barboza, David, "Alibaba Buying South China Morning Post, Aiming to Influence Media", *New York Post*, 11 December 2015; accessible at https://www.nytimes.com/2015/12/12/business/dealbook/alibaba-scmp-south-china-morning-post.html.

39 Redden, Elizabeth "New Scrutiny for Confucius Institute", Inside Higher ED, 26 April 2017; accessible at https://www.insidehighered.com/news/2017/04/26/report-confucius-institutes-finds-no-smoking-guns-enough-concerns-recommend-closure; https://www.nas.org/projects/confucius_institutes; http://www.pewresearch.org/fact-tank/2016/03/30/6-facts-about-how-americans-and-chinese-see-each-other/.

40 Rappler is part of this international research coalition led by Camille François.

41 Nyst, Carly, "Patriotic trolling: How governments endorse hate campaigns against critics", *The Guardian*, 12 July 2017.

42 A definition and discussion of the spiral of silence is available here: https://masscommtheory.com/theory-overviews/spiral-of-silence/.

43 Posetti, Julie, "Online Harassment: Lessons from the Philippines", *Global Investigative Journalism Network*, 13 July 2017; accessible at https://gijn.org/2017/07/13/fighting-online-harassment-lessons-from-the-philippines/.

44 Weedon, Jen, Nuland, William and Stamos, Alex, "Information Operations and Facebook", 27 April 2017; accessible at https://fbnewsroomus.files.wordpress.com/2017/04/facebook-and-information-operations-v1.pdf.

45 Conversation with Facebook, 13 June 2017.

46 Daguno-Bersamina, Kristine, "Duterte uses SONA time to lash out at media", *Philippine Star*, 24 July 2017; accessible at http://www.philstar.com/headlines/2017/07/24/1721155/duterte-uses-sona-time-lash-out-media.

47 Wang, Shan, "Facebook rules the Internet in the Philippines. Rappler walks the line between partnership and criticism", Nieman Journalism Lab, 19 July 2017; accessible at http://www.niemanlab.org/2017/07/facebook-rules-the-internet-in-the-philippines-rappler-walks-the-line-between-partnership-and-criticism/.

48 Romm, Tony, "Governments in 30 countries manipulated media online to silence critics, sow unrest or influence elections", Recode, 14 November 2017; accessible at https://www.rappler.com/technology/news/188536-philippines-freedom-house-freedom-of-net-2017.

49 Simonyan, Margarita, "Russian media from within", Russia Today, 18 October 2011; accessible at https://daily.afisha.ru/archive/gorod/archive/ministry-of-truth-simonyan/.

50 Reuters Institute Digital News Report 2017; accessible at https://reutersinstitute.politics.ox.ac.uk/sites/default/files/Digital%20News%20Report%202017%20web_0.pdf.

APPENDIX A

WORKSHOP AGENDA

WHO SAID WHAT?

THE SECURITY CHALLENGES OF MODERN DISINFORMATION

An unclassified seminar of the Academic Outreach program of the Canadian Security Intelligence Service (CSIS)

20 November 2017, Ottawa

PROGRAM

8:30 – 8:45	Opening remarks: Context and objectives of the seminar
8:45 – 9:30	**Scene-setter** – *Russia, the West and the geopolitics of disinformation: What to expect?*
9:30 – 11:00	**Module 1** – *What is the modern disinformation movement and who are the non-state actors behind it*
11:00 – 11:15	Break
11:15 – 12:15	**Module 2** – *China and the Philippines: Lessons learned and future considerations*
12:15 – 13:15	Lunch
13:15 – 14:45	**Module 3** – *Russia's role in the disinformation movement: Current practise and future prospects*
14:45 – 15:00	Break
15:00 – 16:00	**Module 4** – *The way forward: How to minimise, counter or prevent the impact of disinformation*
16:00 – 16:15	**Synthesis** – *The national security implications of disinformation*
16:15 – 16:30	Closing comments
16:30	Adjourn

APPENDIX B

ACADEMIC OUTREACH AT CSIS

Intelligence in a shifting world

It has become a truism to say that the world today is changing at an ever faster pace. Analysts, commentators, researchers and citizens from all backgrounds—in and outside government—may well recognise the value of this cliché, but most are only beginning to appreciate the very tangible implications of what otherwise remains an abstract statement.

The global security environment, which refers to the various threats to geopolitical, regional and national stability and prosperity, has changed profoundly since the fall of Communism, marking the end of a bipolar world organised around the ambitions of, and military tensions between, the United States and the former USSR. Quickly dispelling the tempting end of history theory of the 1990s, the 2001 terrorist attacks on the United States, as well as subsequent events of a related nature in different countries, have since further affected our understanding of security.

Globalisation, the rapid development of technology and the associated sophistication of information and communications have influenced the work and nature of governments, including intelligence services. In addition to traditional state-to-state conflict, there now exist a wide array of security challenges that cross national boundaries, involve non-state actors and sometimes even non-human factors. Those range from terrorism, illicit networks and global diseases to energy security, international competition for resources, and the security consequences of a deteriorating natural environment globally. The elements of national and global security have therefore grown more complex and increasingly interdependent.

What we do

It is to understand those current and emerging issues that CSIS launched, in September 2008, its academic outreach program. By drawing regularly on knowledge from experts and taking a multidisciplinary, collaborative approach in doing so, the Service plays an active role in fostering a contextual understanding of security issues for the benefit of its own experts, as well as the researchers

and specialists we engage. Our activities aim to shed light on current security issues, to develop a long-term view of various security trends and problems, to challenge our own assumptions and cultural bias, as well as to sharpen our research and analytical capacities.

To do so, we aim to:

- Tap into networks of experts from various disciplines and sectors, including government, think-tanks, research institutes, universities, private business and non-governmental organisations (NGOs) in Canada and abroad. Where those networks do not exist, we may create them in partnership with various organisations;
- Stimulate the study of issues related to Canadian security and the country's security and intelligence apparatus, while contributing to an informed public discussion about the history, function and future of intelligence in Canada.

The Service's academic outreach program resorts to a number of vehicles. It supports, designs, plans and/or hosts several activities, including conferences, seminars, presentations and round-table discussions. It also contributes actively to the development of the Global Futures Forum, a multinational security and intelligence community which it has supported since 2005.

While the academic outreach program does not take positions on particular issues, the results of some of its activities are released on the CSIS web site (http://www.csis-scrs.gc.ca). By publicising the ideas emerging from its activities, the program seeks to stimulate debate and encourage the flow of views and perspectives between the Service, organisations and individual thinkers.