CHAPTER 2

Propaganda and Disinformation Go Online

Miroslava Pavlíková, Barbora Šenkýřová, and Jakub Drmola

2.1 Introduction

What is the future of political propaganda? Is the future 'now'? In Ukraine, information warfare was made public, and a discussion about new forms of propaganda dissemination subsequently appeared. Human controlled accounts manipulating online content at the same time emerged in the Russian–Ukrainian conflict. The so-called Kremlin trolls tried to influence domestic as well as Ukrainian audiences in favour of the Russian framing of the confrontation. The Ukrainian side, for its part, also attempted to utilise the Internet as a non-kinetic weapon. The Ukraine Information Army was founded by the Ministry of Information Policy as an initiative aiming to fight the Russian trolls, and every Ukrainian citizen

M. Pavlíková (✉) · B. Šenkýřová · J. Drmola
Department of Political Science, Faculty of Social Studies, Masaryk University, Brno, Czech Republic

J. Drmola
e-mail: jdrmola@mail.muni.cz

M. Gregor and P. Mlejnková (eds.), *Challenging Online Propaganda and Disinformation in the 21st Century*, Political Campaigning and Communication,
https://doi.org/10.1007/978-3-030-58624-9_2

was able to become a member of the group, a tactic aimed at debunking the disinformation spread by the trolls.

Another sign of the times as regards technology development and propaganda usage was the 2016 US presidential election and the ensuing allegations about Russian state interference. Russia conducted complex influence operations with the use of cyberspace, where propaganda played a significant role. A massive campaign of robotised accounts spreading content hit social media. Precisely chosen target groups became the objective of this upgrade to the Ukrainian Internet's former Kremlin trolls. Shortly thereafter, a scandal with the company Cambridge Analytica appeared. After the Brexit referendum, it was revealed that the consulting company had misused holes in the data protection of Facebook users to produce micro-targeted political campaigns all around the world—the company's services had been hired by state actors on various levels. What is more, we are now reading about the threat of the so-called deep-fakes and their ability to perfectly deceive a target audience. Could future propaganda and information manipulation like these examples decide elections?

A new information environment, driven largely by the growth in the Internet, is rapidly changing the economic, social, and political landscape. Given the exponential growth in both Internet use and the availability of news, it is no wonder this new form of communication brings new tools and methods for media manipulation. The aim of this chapter is therefore to introduce and describe changes in the information environment and the newest tools and trends identified in the context of information manipulation and propaganda at the beginning of the twenty-first century. The technology of disinformation and propaganda has been going through significant changes since the end of the Cold War. Together with changes in society and the very meaning of information, we could nowadays talk about an information society (Webster 2006), information economy (a society that produces more information then goods) (Zhang 2017), or an information-centric society. The societal changes have also applied to the security and military domain. Although the importance of information dates to classic military theorists like Sun Tzu, its connection with the rapid emergence of new technologies has profoundly changed the way conflict and war is waged. Terms like information warfare or the weaponisation of information (see Toffler and Toffler 1981) have become a vital part of today's security and military discourse.

Sun Tzu would agree that information has always been an extremely dangerous weapon. However, at the start of the twenty-first century, we are witnessing very turbulent times due to high-speed technological development and the expansionary potential of virtual space. The chapter therefore begins with a reflection on the information environment and the infosphere. The focus is put on the development of the information environment, because that is the decisive factor affecting what tools can be used and in what way information can be manipulated. The issues related to the information environment and the efforts of countries to protect it is illustrated by the cases of two undemocratic countries, China and Russia. These countries are more prone to believe disinformation and propaganda to be suitable parts of a toolset in the protection and promotion of their interests, and efforts to control the information environment go hand in hand with it. Through such examples we can demonstrate the importance of the information environment (and its protection), and also how far regimes can go in their obsession to control the flow of information given the current circumstances. How different information environments can work as well as how they can prevent the public from being informed within a regime will be described, but also, conversely, how they can openly spread information outside a regime.

The second part of the chapter is focused on technological disinformation and propaganda tools in online space. This part depicts the shift from human-made trolling to bot activities and highly personalised, precisely targeted, emotion-adaptive artificial intelligence. The following section briefly confirms that it is not only actors with political or ideological goals using disinformation and propaganda—a frequent motive for deploying them is in fact profit. Another change observed in the context of online disinformation and propaganda is that of businesses implementing manipulative activities in political communication, as can be seen in the example of the 2016 US presidential elections and Cambridge Analytica; and let us not also forget the crucial role social media platforms, such as Facebook, played as well. Technology companies will therefore be subsequently covered. In the final part, we move closer to the future, focusing on deepfake technology, which can be seen as a future tool of new online propaganda. Although deepfakes are a very new disinformation technology, we can already explore it through some examples (fake videos of Barack Obama and Donald Trump, fake profile pictures on LinkedIn, and the FaceApp data privacy case). Although deepfakes represent new

technology, we argue that they mark the next step in the gradual professionalisation of information manipulation rather than a revolutionary new phenomenon.

The aim of this chapter is not to provide readers with an exhaustive list of 'high-tech technologies' but rather to point out the main directions of modern propaganda and disinformation developments.

2.2 Information Environment

The key element all propaganda and disinformation are based on is information. Since the development of technology such as radio and television, the importance of information and its value has been changing. In the 1950s, when television broadcasts entered the daily lives of people in the developed world, the perception and need for new information began to be a new requirement in their lives. The word information (having Latin roots and a Greek origin) is used in two basic contexts: the act of moulding the mind and the act of communicating knowledge (Capurro and Hjørland 2005, 351). The action of *informing*, therefore, refers to the shaping of mind or character, education, and teaching (Oxford English Dictionary). According to the OECD there are three conceptions of information: information as knowledge, information as economic activity, and information as technology (Godin 2008, 256). In our chapter we will deal with all of these conceptions. As a society in which information and information technology have a powerful influence on everyday social life (OECD, *A Framework document in* Godin 2008, 268), we often use the term 'information society'. Similar to information itself, this concept is not clear and generally accepted. It is worth pointing out here that one of the approaches to defining information society is the fulfilment of these five criteria: technological (connected with technology innovation since 1970), economic (in the context of the information economy), occupational (informational work), spatial (change in the organisation of time and space), and cultural (information in everyday life—TV, emails, etc.) (Webster 2006, 8).[1]

The definition of information as it relates to the information economy is 'grounded more in the production and exchange of information than

[1] According to Webster (2006, 8), there is also a sixth definition. However, '… its main claim is not that there is more information today (there obviously is), but rather that the character of information is such as to have transformed how we live'.

physical goods' (Mazarr et al. 2019, 3). Moreover, in some studies, information society is used instead of information economy (Atik 1999, 121). According to OECD information economy 'refers to the implications of information technologies on the economy, [and] on firms' performance (productivity, profitability, employment)' (Godin 2008, 268). Godin's three conceptions here demonstrate that information and information technologies are, today, the main commodities of international business.

The information environment, where all these processes take place, is both a vaguely and complexly understood concept. Media and communication studies often use terms like 'press system', 'media system', and 'mass communication system'. However, the way in which media systems are constructed and developed is not homogenous (anymore). Moreover, these concepts do not integrate all the actors and ways of interpersonal communication, which we can see today. Therefore, the concept of information environment provides a better understanding of the 'battleground' where propaganda and disinformation fight for our attention and affection. The information environment consists of three dimensions: cognitive, informational, and physical (US Joint Chiefs of Staff 2006). The information environment throughout most of the twentieth century was one in which people could easily identify the place, time, and space information was related to; however, this became more difficult (if not impossible) with the advent of the Internet in the late 1980s. This shift not only made it difficult for consumers to identify where and when information was initially released, but even identifying the producer became challenging. Another way of conceiving this change is that cyberspace and social networks became intertwined with (or started overlapping) the information environment (Porche et al. 2013, 12).

The integral part of the information environment is the infosphere. Sometimes, these two terms are used synonymously. However, in our view, these two concepts are more complex, and, therefore, we will define the infosphere separately. The term was devised by Boulding in 1970. He notes that 'the infosphere then consists of inputs and outputs of conversation, books, television, radio, speeches, church services, classes, and lectures as well as information received from the physical world by personal observation. … It is clearly a segment of the sociosphere in its own right, and indeed it has considerable claim to dominate the other segments. It can be argued that development of any kind is essentially a learning process and that it is primarily dependent on a network of

information flows' (Martens 2015, 332). Similar to the information environment, the infosphere used to be controlled by only a few television channels and shows (Mazarr et al. 2019, 25). Regularly published newspapers and the continuous broadcasting of radio and TV stations were supplemented by the Internet, where access to the information is more a question of the users' will, deciding for themselves when and where they wish to consume it. In these terms, and given the permanently changing environment and society, infosphere could be defined as 'the ongoing process of producing, disseminating, and perceiving information in a society, including media, data-based algorithmic processes, and information exchange in networks' (Mazarr et al. 2019, 7). In other words, in today's information environment, 'the direct links between time and place, on the one hand, and the individual as a producer and consumer of the content, on the other hand, have long since disappeared' (Brikše 2006). This shift has increased the space and possibilities for the production of media manipulation. This also applies to the infosphere, whose challenges we will discuss later.

In the context of today's information environment, cyberspace, which is essential but difficult to define, is often mentioned. The word cyberspace was first used by writer William Gibson in 1982 in the short story 'Burning Chrome'. In the absence of a widely accepted definition of cyberspace, each sector defines the term in accordance with its own needs. In recent years, however, there has been a trend to call everything related to networks and computers 'cybernetic' (Lorents and Ottis 2010). In addition to the Internet and computer networks, we also include telecommunication networks, embedded processors, and drivers in cyberspace. For example, Lorents and Ottis (2010) attempted to create a general definition of cyberspace as a virtual world which would be valid in all sectors. According to them, cyberspace represents a time-dependent set of interconnected information systems and human users interacting with these systems. Although cyberspace is independent of geographical boundaries, it falls within the jurisdiction of a state located in that territory. However, even this division is unclear and legally difficult to grasp, as will be shown later in this chapter. At the same time, it makes it challenging to enforce legal offences and crimes committed by individual users in cyberspace.

State and non-state actors endeavour to protect their cyberspace. These efforts, called cybersecurity, can be represented by the concepts of IT security and electronic information system security. There are de facto two approaches to cybersecurity as a discipline: the technical and social

science approaches. The first deals mainly with the practical side—the actual repelling of cyber attacks, their analysis, the implementation of technically oriented security solutions, and more, i.e. the logical and physical layer of cyberspace—while the latter explores the social layer of the problem.

Because information technology is hugely dynamic, attackers are continually inventing new techniques of malicious activity. Those that were up to date five years ago appear to be outdated today and have been replaced by more sophisticated methods. This development in cyberspace requires flexible responses from the security community. Knowing the source of the attacks is more important than knowing the attack technique. Although cyber threat techniques change over time, the actors usually remain the same (Secureworks 2017).

2.3 Challenges in Today's Information Environment

As the technology developed (through the spread of different broadcast methods), so too did the infosphere change. Nowadays, the infosphere could be characterised by many (not only) new problematic trends, including 'the fragmentation of authority, the rise of silos of belief, and a persistent trolling ethic of cynical and aggressive harassment in the name of an amorphous social dissent' (Mazarr et al. 2019, 2). It means that today the infosphere is determined by:

- networked dynamics and the role of viral spread of information
- broad-based sensationalism in news and other media
- fragmentation of the infosphere
- concentration of information platforms
- the effect of self-reinforcing echo chambers
- the role of influencers
- the emergence of a 'trolling ethic' on the Internet
- the explosive growth of data collection on individuals and groups. (Mazarr et al. 2019, 19)

All of these characteristics influence how the information environment is perceived and could indicate the main future challenges (largely in the context of online propaganda, cyber attacks, and information warfare). In what follows, we will focus on examples of how state or non-state actors

react to changes in the infosphere by controlling the information environment. However, the main challenges will be the new tools used in the current information environment. This approach is called hostile social manipulation, and it uses techniques like trolling or deepfake learning.

The Internet, its development, and its expansion is generally tied to freedom of speech (see Chapter 4) and the process of democratisation (Whitehead 2002; Grugel and Bishop 2013). Take, for example, the significant role it played in the events surrounding the Arab Spring (see Wolfsfeld et al. 2013; Khondker 2011). However, together with Internet development and information dissemination, the tendency towards information control has grown as well. Nondemocratic actors appear nervous due to the flow of information, which is almost impossible to stop given the ease of accessibility. They are afraid of an erosion in their political power and, therefore, search for methods of controlling the situation as much as possible—via censorship or manipulation. An example of these tendencies is the contemporary absence of visible political changes in authoritarian regimes, which indicates that Internet censorship helps these regimes consolidate (Tang and Huhe 2020, 143).

Such a situation is evident in China, where the Internet is censored, monopolised by the state, and filled with manipulative content (see Freedom House Report: China, 2018). Starting in 1997, when the Internet began to be widely used in China, the government passed laws and regulations providing control over ownership connected with all Internet services and media, Internet content, and all behaviour on the Internet (Lu and Zhao 2018, 3297; Thim 2017). Later, in 2014, a new Office of the Central Leading Group for Cyberspace Affairs (CAC), controlled by the General Secretary of the Communist Party of China, was established. This office controls the regulation of the Internet with its justification being cybersecurity (Lu and Zhao 2018, 3297). In 2016, China passed the so-called Cybersecurity Law, which strengthens the role of the state in cyberspace and emphasises security over freedom of speech (Qi et al. 2018).

Internet regulation in China is layered in three levels (King et al. 2013, 3). The first is the Great Firewall of China, meaning the specific information environment in which only some websites are allowed. In fact, it means most foreign websites are blocked. This censorship does not allow people to be in connection with people or media from outside of China. The second type is keyword blocking, which makes publishing a text (whether on websites or social media) impossible when a user

writes banned words or phrases. The third level can be called 'hand censoring', and, unlike the previous two types, this one is employed after the publication of information. While the first two methods of censorship are automatic or technical, the latter is carried out by people; there are thousands of censors reading, marking, and prohibiting content on the Internet. A special type of hand censor is people focusing on content production. One of the known examples is the so-called 50 Cent Army of Chinese Internet trolls. They are paid by the government to write comments favouring the Chinese communist regime (Farrell 2016).

Despite its geographical distance, China's operation within the information environment can constitute a threat to the Euro-Atlantic space. Today, discussions about new, fifth-generation information infrastructure (5G) and its provision in Europe and the United States are occurring. The affair became prominent in 2019 when a Pentagon report dedicated to 5G infrastructure mentioned the security risk connected with the Chinese intelligence community (mainly the company Huawei). Based on this report, Huawei was placed on an entity list which prohibited American companies from selling their goods to them (without special government permission). The risks resulting from a 5G network were discussed by many governments in Europe as well—especially in the context of national security. Since Huawei is tied to the Chinese regime, there is concern about its approach to the information environment (see Chapter 7 for the example of the Czech Republic). The distrust originates mostly from the strong connection between the Chinese government (especially secret services) and technology providers such as Huawei or ZET. There is apprehension over China's potential use of the 5G network for influence operations. However, despite evident pressure, Huawei still cooperates and has signed contracts to supply 5G in many countries (Segev et al. 2019).

Besides China, Internet censorship and content adjustments also occur in Syria, India (Steen-Thornhammar 2012, 227–228), North Korea, Cuba (Cheng et al. 2010), and Russia. Cuba and North Korea use the mosquito-net model whereby the governments support the inflow of foreign investment (even in the IT business), but they block the inflow of foreign values, ideas, and culture (Cheng et al. 2010, 660). In North Korea, access to the Internet is limited and available only to elites (Cheng et al. 2010, 650).

The Russian approach is characterised instead by censorship and intimidation rather than limited access (Maréchal 2017, 31). However, Russia

does have its liberal loopholes—the development of RuNet, the Russian-language community on the Internet, was from the beginning mainly about academic sharing and development (Bowles 2006). Nonetheless, it still attempts to widen its Internet environment surveillance. The new Russian information policy is notably connected with the start of the 2010s and the accumulation of power into Vladimir Putin's hands.

State-controlled communication providers have a significant role in Russia, which allows the state, through its secret services, to have easy access to Internet traffic (Freedom House Report: Russia, 2014). The legislature is also a powerful tool in Internet regulation (see legislative measure No. 428884-6, the so-called Bloggers Law; legislative measure No. 553424-6 about data storage; or a legislative measure No. 89417-6 2012 concerning children's protection against malicious Internet content). All these legal measures have attempted or enabled the state to better control Internet content, its regulation, and repression against dissent. Moreover, the laws are flexible in their application and can target a wide spectrum of subjects unfavourable to the regime.

The specificity of the emerging form of RuNet lies in its strong ties with the government and its relation to the Russian national identity. RuNet has been defined as 'a totality of information, communications, and activities which occur on the Internet, mostly in the Russian language, no matter where resources and users are physically located, and which are somehow linked to Russian culture and Russian cultural identity' (Gorny 2009 in Ristolainen 2017, 118). The idea of RuNet and the isolation of the Internet in the Russian context is connected with the idea of Russian sovereign space and challenges to the 'US dominated world' (Ristolainen 2017, 124). Tensions between Russia and the West (represented especially by NATO and the United States) started to deepen, with a peak during the Ukraine crisis and the 2016 US presidential election interference. On 12 May 2019, a bill was adopted which confirmed the control and isolation of the Russian language Internet space from the rest of the Internet as of 2020 (The Moscow Times 2019).

The Russian approach to Internet regulation and online propaganda can be described through the concept of hybrid warfare. Russia uses advantages in the information environment because the low price and simplicity make it available. Instead of regular warfare, it builds on irregularity to counterweigh possible asymmetry in confrontation.

Control over the information environment has multiple benefits for undemocratic regimes. First of all, there are the economic reasons. The

Internet and its development are crucial for the economic development of a country, which might be the main legitimating factor for authoritarian countries (Kalathil and Boas 2003, 144). However, for modern strategic communication, in order to control the minds and opinions of citizens and, therefore, support for the regime's fundamentals, it is important to control and censor the Internet. As examples from Russia and China show, it is easier to establish a sovereign information environment for all of these variables. Jiang (2010, 72–73) offers a theoretical explanation called 'authoritarian informationalism'. It is based on Internet development and regulatory model. Individual responsibilities are emphasised over individual rights; maximum economic benefit and minimal political risk for the one-party state are also stressed. Jiang explains the concept through the Chinese case. Authoritarian informationalism in China combines elements of capitalism, authoritarianism, and Confucianism (Jiang 2010, 82). Jiang further claims that authoritarian informationalism describes the future reality of its information environment because it is based not only on extending control but also on enhancing its legitimacy (mainly based on trust in government and economic success).

These characteristics pertaining especially to authoritarian countries have produced new threats for the democratic world as well as their own civilians. The authoritarian regimes tend to control their own information environment inside the country and, at the same time, interfere in the outside environment so as to reduce the power of information which might be endangering the regime and running against its interests. In the following sections, the latest technological tools and tactics of propaganda and manipulation will be approached. They are the tools used to control and manipulate the information environment. What looked like science fiction in the 1970s has become a reality today or will in a few years' time.

2.4 From Trolls to Bots

In 2015, there was an article called 'The Agency' published in the *New York Times* (Chen 2015). This occurrence could mark the moment interest began in the new wave of Internet propaganda, especially for the Western world. The article by Chen presented the wider public with the existence of the so-called Kremlin trolls and their organisation. 'The Agency', believed to be linked to Putin's administration, refers to one

of the first publicly known 'troll farms' (*Financial Times* 2019). This one, located in St. Petersburg, Russia, was employing mostly young Russians who were uninterested in politics. As part of their job description, they were supposed to spread propaganda regarding the ongoing Russia–Ukrainian conflict without any personal interest.

This agency represents a new tool of propaganda strictly linked to the online world: *trolls*. With advantages for propagandists and disadvantages for receivers, trolls, sometimes referred to as hybrid trolls (Hannah 2018), are persons who aim to spread or destroy a particular narrative. Their primary role is to dominate the discussion on social media or in discussion forums; to overwhelm it with various contributions, often not even related to the topic of discussion; and to vulgarly offend opponents' opinions and discourage them from further discussion using these practices (for a larger explanation of the term and its role in cyberspace, see Nycyk 2017). In this case, trolls are acting as a force multiplier for driving home Russian messages (Giles 2016, 9). As leaked documents proved, paid trolls from St. Petersburg received worksheets every day with indications of the topics they should cover and discourse they should use.

The story of trolls from St. Petersburg is connected to an inconspicuous administrative building in the area of Olgino where mostly young people without any significant political stance work basically as copywriters. However, the content they were producing was strictly oriented to the framing of the ongoing political situation in Russia and Ukraine. This is in marked contrast to what would be seen three years later when trolls with the same background conducted sophisticated complex influence operations, including classical trolling, fake website making, local news outlet impersonation, cooperation with Russia's military service (StopFake 2019), and the production of micro-targeted campaigns.

Besides the Ukraine conflict, the troll's operations were spotted in the Baltics a few months later—according to research from the NATO StratCom Centre of Excellence, the activities of the St. Petersburg troll farm were evident in Latvia (Spruds et al. 2016). Finnish investigative journalist Jessikka Aro also revealed[2] Kremlin troll activities in Finland, which helped to publicise the issue; however, her investigation also resulted in massive cyberbullying directed towards her (Aro 2016; Rose 2019).

[2] Aro's case went even further. The most aggressive trolls/cyberbullies ended up in Finnish court, and three people were consequently sentenced (Staudenmaier 2018).

What might have seemed like regional information operations evolved into a supranational conflict and propaganda campaign. St. Petersburg's trolls became known by its official business name, the Internet Research Agency (IRA). The IRA was found to have interfered in the 2016 US presidential election (US Department of Justice 2019), but Howard et al. (2018) noticed that the targeting of US society by the IRA had begun much earlier. Engagement of IRA trolls was further proved in the 2016 Brexit referendum campaign (Bastos and Mercea 2018; Field and Wright 2018) as well as in the 2017 German general election and the debate which followed a mosque shooting in Canada (Al-Rawi and Jiwani 2019). The Kremlin trolls' interference in the US election demonstrated just how sophisticated this form of online propaganda could be, and how its operational abilities, scope, and intensity are growing (Inkster 2016; Badawy et al. 2019; Kreiss 2019).

The trolls had been covering a wide range of topics, mostly with the potential to divide public society. To illustrate, they massively covered topics connected to black American culture as well as veterans' issues and gun rights in the US election campaign. Targeted groups were sometimes contradictory, such as covering anti-refugee as well as pro-immigration reform content, yet they were deliberately selected with a focus on message receptivity (Yonder 2018; Howard et al. 2018). The case illustrates one of the main characteristics of these online information operations: the sowing of uncertainty, chaos, and the fragmentation and polarisation of society, as opposed to the propagation of a particular ideology.

As demonstrated above, the issue of trolls is closely tied to pro-Kremlin individuals and groups in European discussions, because these are the most active propaganda actors influencing the Euro-Atlantic region. However, we can identify many other organised groups of trolls all around the world, for example, the above-mentioned Chinese 50 Cent Army producing content for its own citizens to support the regime (see Farrell 2016). During the Brexit referendum, it was not only Russian trolls who interfered, Iranian trolls operating on Twitter took part as well (Field and Wright 2018). In 2018, Twitter and Facebook shut down hundreds of fake accounts of Iranian government origin. The trolls promoted Iranian political interests focusing on anti-Saudi, anti-Israeli, and pro-Palestinian themes as well as topics targeting US politics (Titcomb 2018). Trolling groups have even emerged in Europe. One example is the extreme-right group Reconquista Germanica, focusing on German politics and

sympathising with the far-right party Alternative for Germany (Ebner 2018).

The emergence of troll campaigns was not expected, and, in the beginning, shared content could have been easily confused with real users blogging. Yet, while the phenomena of online troll activities have been recently unveiled, their tactics have already developed, and their sophistication is growing. The evolution of trolling has not been restricted to only geographic expansion and better targeting, its robotisation and automatisation is also taking place.

Automated propaganda or robotic propaganda is based on the activities of *bots*—programmes automatically producing content that should look like that of real users, interact with humans online, and produce manipulative content on social media, especially on Twitter, Facebook, and Instagram (see The Computational Propaganda Project 2016; Gorwa and Guilbeault 2018; Nimmo 2019). Originally, they were used as a supplementary tool for trolls, with bots spreading the content produced by trolls and genuine pro-government users.

Low costs, availability, and scaling through automatisation (Woolley and Howard 2016, 7) have shown that bots alone can help governments, political parties, or other interest groups to manipulate an audience's opinions. Naturally, the main difference between trolls and bots lies in bots' ability to coordinate tweeting about the same issue thousands upon thousands of times a day (see Gorwa and Guilbeault 2018, 10). In many examples, the programmers who deploy bots work as pure mercenaries. They are ideologically apolitical and motivated strictly by money (Woolley and Howard 2016, 10).

By using online bots, propaganda can produce the impression of a strong grassroots campaign and, therefore, attract new supporters or encourage higher activity among existing ones—a practice called astroturfing (see Zhang et al. 2013; Spruds et al. 2016; Kollanyi et al. 2016). Twitter represents an ideal platform for this strategy; a large amount of bots and real users were actively in favour of Donald Trump's and Hillary Clinton's Twitter campaigns (with hashtags '#MAGA' and '#ImWithHer') as well as the Brexit referendum. Through a high frequency of tweeting and coordinated activity, bots are able to shift and distort 'trending' topics, messages, and posts with hashtags. It is usual for bots to 'hijack' hashtags. Popular and trending hashtags are exploited with the intention of getting more visibility and attention even though the content shared by bots does not have to be connected to

the topic of a hashtag anyhow (Woolley and Guilbeault 2017). Take, for example, the coordinated hijacking of Twitter hashtags during the 2017 Florida high school shooting. Trending hashtags like #NRA, #shooting, #Nikolas, #Florida, and #teacher were hijacked by political bots, mostly originating in Russia (Frenkel and Wakabayashi 2018).

Bots are also reusable. Sleeping botnets, a system of bots, can be inactive or focused on non-political spamming for sometime. When it is needed, they can be activated to spread political and manipulative content. Therefore, one group of bots can be used during multiple political campaigns by different actors with various ideologies (Neudert 2017). Besides that mentioned above, bots are deployed as part of various strategies to manipulate opinion. They can also maintain coordinated intimidation, participate in surveillance against citizens or regime opponents (Pavlíková and Mareš 2018), help block certain content through coordinated complaints to providers, or be used as a tool for search engine optimisation (SEO) (Zhdanova and Orlova 2019, 53–54).

Troll and bot activities spreading particular narratives should not all be perceived as strictly directed by governments. The specificity of contemporary online propaganda is its participatory character (see Wanless and Berk 2017); there is, besides attempts to manipulate audiences, the co-option of members from this audience to active engagement in propaganda too. Manipulation by trolls and bots is also used by political parties or individual politicians, particularly during election campaigns.

Bot activities can be massively enhanced by new technologies based on artificial intelligence (AI). The Atlantic Council defines this phenomenon as the integration of artificial intelligence into machine-driven communication tools for use in propaganda—MADCOM (Chessen 2017). MADCOM uses machine learning, deep learning, and chatbots. When machine learning is combined with big data, a very powerful propaganda tool emerges (Chessen 2017). Campaigns become highly personalised based on information about recipients' activities in virtual space and information shared in virtual space about family, friends, political preferences, demographic data, and hobbies and, therefore, precisely targeting people's individual characters and, what is more, precisely targeting vulnerabilities and detecting emotions in real time. These sophisticated technologies, when linked to private companies willing to be hired by governments, parties, or even individuals, may bring radical inputs into decision-making processes or voter behaviour.

2.5 Blurring the Lines Between Politics and Business

In the information society, boundaries between politics and business are being blurred. The line between public relations and propaganda has always been rather fuzzy and their development has been closely connected for at least a century (see Bernays 1928). In this chapter we will introduce how the strategies and technologies more recently developed by the business sector are now also used by political manipulators while, simultaneously, propaganda has become a profitable business. This can be considered a typical trait of an information society and economy, as detailed earlier in this chapter, in which information is the primary business commodity and source of power and political influence. In this sense, just as the information and information technologies became valuable for business, they also became valuable to political competition. At present, developed countries face a problematic lack of political programmes and political values, e.g. populism is on the rise. At the same time, traditional means of political competition (such as distinct values and ideologies) are today becoming increasingly supplemented (or even displaced in some cases, one might argue) by business-powered influence campaigns. Crucially, technological proliferation and the resultant data accumulation is enabling these efforts, both in terms of their large scale and also precise targeting.

The Internet Research Agency (IRA), originally a hybrid between a private subject and a state-controlled organisation, once again provides an example. At some point, it gradually shifted from the use of trolls to more precise information campaigns based on accurate group targeting, which has been more typical for business. As analysis of the US election interference shows, the IRA was thoughtfully focusing on segments of social media users based on race, ethnicity, and geographical division, and it ran multiple ad campaigns targeting different groups (Howard et al. 2019). Its strategy had two stages. The first was focused on the narratives of a specific group as a clickbait strategy to drive traffic to the IRA's pages. The second was to manipulate the audience by posting content to these pages (Howard et al. 2019, 18–19). The case of the IRA emphasises complexity and level of online manipulation threats which lie at the border between state propaganda and the business model, merging both and learning from each.

There is also a growing trend today of cyber troops being deployed primarily as a tool of mass (yet targeted) influence. The character of bots, their reusability and effectivity, has led to the formation of IT propaganda mercenaries. Woolley and Howard (2018, 10) mention programmers who deploy bots for hire and who are purely money-oriented, without any political affiliation. These can be equally deployed to influence potential customers as much as potential voters.

Besides cyber mercenaries, there are whole companies participating in online political manipulation (without connection to a government as it was with the IRA). An example from the Brexit campaign shows that big companies can be powerful in opinion-forming and can lead to huge political consequences. New technologies using micro-targeting, machine learning, deep learning, and big data—mostly from social media providers with weak user data security—are able to manipulate opinions to advance ideological viewpoints (Neudert 2017). Connections between campaigners for leaving the European Union and the companies Cambridge Analytica and Aggregate IQ doing 'psychological warfare' (Cadwalladr and Townsend 2018) formed the opinions of Britons during the Brexit referendum.

Cambridge Analytica promoted itself as a hi-tech consultant using data to micro-target population groups with precisely designed messages. However, according to accusations, the company harvested the data of 50 million Facebook users to set up the campaign (Cadwalladr and Graham-Harrison 2018). Christopher Wylie, a former employee and whistle-blower, later stated that the company 'exploited Facebook to harvest millions of people's profiles. And built models to exploit what [Cambridge Analytica] knew about them and target their inner demons' (ibid.). While political actors have been hiring private businesses for public relations for decades, these micro-targeting propaganda campaigns, which have proven especially effective at manipulating a targeted audience, have only been enabled by for-profit technologies deployed exclusively by business actors since they require some form of access to massive amounts of private customer and/or user data in order to operate.

Cambridge Analytica did not operate only in the United Kingdom but in many other countries during important elections, mostly in Africa or the United States (Solomon 2018). However, the example of the Brexit referendum campaign and its consequences emphasises the scope of the overlap between politics and business. Moreover, this internal form of propaganda took place in a democratic state. It emphasises that two

cherished attributes of democracies—freedom of speech and economic freedom—can also make them more vulnerable.[3]

2.6 Deepfake as a Future Tool of Online Propaganda?

The Internet offers a platform to anonymously spread manipulative content which—through use of new dissemination technologies—is so easy that anybody is able to produce it. Depersonalised and anonymous accounts can enable and amplify their hostile messaging. What has always been challenging for manipulation was the credibility and trustworthiness of the information. Nevertheless, with new forms of manipulation techniques and new technologies, including artificial intelligence, this issue seems likely to be overcome very soon. In 2017, a new tool usable for propaganda was made available: the so-called 'deepfake' (see Parkin 2019). The reason why this new phenomenon came to be studied was the introduction of new deep learning technology. As one of its uses, this technology could alter and replace the faces of people in videos with believable results. Even more importantly, it is no longer an exclusive tool of IT engineers and enthusiasts, but a tool that is becoming available to the wider public.

The deepfake is a combination of two expressions: *deep learning* (a product of artificial intelligence) and *fake news* (false content, see Chapter 1). The term refers to false audiovisual content generated by artificial intelligence which is so credible that the average viewer is unable to recognise it as a product of artificial intelligence or falsified content (Chesney and Citron 2018a). Software using modern deep learning technology combs through a large amount of data (videos, images, or audio tracks), analyses them, and over time learns to recognise regularities, such as intonation, facial expressions, voice colour, or gestures. These AI systems have two elements: generator and discriminator. The basic function of the generator might be to create a new video—a fake video clip. Thereafter, the discriminator is asked to test if the video is real or fake. Based on the discriminator's evaluation, which identifies what the fake parts are, the generator learns what to avoid in the next

[3] Naturally, online internal state propaganda in nondemocratic regimes takes different forms. Botsman (in: Jiang and King-wa 2018) describes Beijing's propaganda as 'Big data meets Big brother' (plus Big profit).

video clip. When a generator produces an acceptable level of output, the videos are again 'fed' to the discriminator. The process is repeated so the generator gets better at creating videos and the discriminator gets better at analysing them. These two parts of the system are also called a generative adversarial network (GAN) (Yuzeun and Lyu 2018, 47; Robinson 2018, 18; What Is It 2019, n.d.; Giles et al. 2019).

While discussing deepfakes, many people might be familiar with the 'face swap', as an example, when the face of another person who does not initially appear in a video is added into the scene replacing someone else. It is a simpler method of creating this type of fake video, and it is widely used in pornographic content. However, what is possible by using deep learning technology and what is correctly indicated as a deepfake is the generation of custom content based on the analysis of accumulated data (as opposed to simply swapping two pieces of pre-existing content). This type is not only more sophisticated but also represents a greater security threat because, with sufficient data, it can realistically simulate any expression of any person (e.g. influential politicians).

The first huge scandal connected with deepfake technology was the distribution of pornographic videos with faces of celebrities like Gal Gadot or Emma Watson, who were not involved in the production (Chesney and Citron 2018b; What Is It 2019, n.d.). The fake or manipulated videos and photoshopped stills are there for decades and more are being produced all the time. With the expansion of social media, which provides perfect source material to train the algorithms, there are numerous examples of fake photos and fake videos which could affect the security and well-being of celebrities and citizens alike.

Among the most famous examples of a fake video (which started the public discussion about deepfake) is a 'phone call' between Barack Obama and Donald Trump. In 2016, NBC's *The Tonight Show* aired a scene featuring Jimmy Fallon dressed up as Donald Trump and on a phone call with Dion Flynn, who was dressed up as Barack Obama. This scene was remade in 2019 when a similar video was uploaded onto YouTube by a user called James. This time, however, the video was created through the deep machine learning model and viewers had the impression of watching the real faces of Donald Trump and Barack Obama. Their gestures and voices were almost indistinguishable from those of the real presidents (Parkin 2019). Along with the increasing quality of deepfake videos comes a growing concern that the abuse of such a tool will lead to attacks on

politicians, which could be especially harmful during elections. Danger arises from the fact that today, there is no readily available software which could quickly and reliably detect these videos as fake (Schellmann 2017).

However, deepfake technologies are not just videos. Another use of the technology was demonstrated in the case of a fake Kate Jones LinkedIn profile photo. Jones' profile differed from 'ordinary' fakes in many ways—style, precision of the profile, but mainly in the use of deepfake technology. Instead of a copy of a photograph or a stolen photograph from a real person, Jones' fake profile used a unique photo, which was a computer-generated artefact made by machine learning algorithms. This made it more difficult for real LinkedIn users to recognise that the account was fake and many of them accepted the connection requests from Jones. Requests were accepted even by the US Defence Attaché to Moscow, a top-ranking US State Departmental official,[4] and other professionals (Giles et al. 2019, 4–5). The profile was detected as fake in June 2019, three months after its creation. The purpose of the fake profile as well as its creators remain unknown. This 'faked authenticity' is a cause for concern, as the erosion of standard and easily accessible methods for verifying accounts (such as reverse-searching the photos) makes distinguishing fake profiles from real ones more difficult.

Another tool connected with deepfake threats in 2019 was the application FaceApp. The app is based on AI algorithms that can swap faces in videos. FaceApp shows that using tools employing artificial intelligence can be so simple that anyone can do so using a smartphone. It does not need any sophisticated hardware systems or a team of specialists and experts. In addition to funny videos, the app can be used to produce fake videos (Dickson 2018). Another threat FaceApp presents lies in the data privacy field. In 2019, when celebrities posted their photos generated by this app on social media, starting the 'FaceApp challenge', the app became famous and widespread among dozens of millions of users. As it turned out later, the company responsible for the app came from Russia, and the app's privacy policy is unclear; users do not know if the company is collecting their data or selling it to third parties, or possibly harvesting their data to train even better deep learning algorithms for future deployment. Moreover, this type of application could be a tool for

[4] Authors of the research did not mention the names of officers.

the future politically oriented goals of some organisations (see Wondracz 2019; Libby 2019). With the company based in Russia and in the context of information warfare, the case is more relevant—a US investigation into the app and its data handling was started in July 2019 (Wondracz 2019).

Deepfake technology is also not focused exclusively on visual content. A necessary part of any (video) content is the sound as well, and it is even easier to manipulate audio than video. Two of the main projects dealing with deepfake technology in the audio environment include 'Project VoCo' created by Adobe, which is nicknamed 'Photoshop for Audio' (Wardle and Derakhshan 2017, 76), and the second is the widespread tool Lyrebird, which is focused on deep learning and neural networks and is used to synthesise the human voice (Dickson 2018). Lyrebird represents another example of an easy-to-use app; it needs just a one-minute recording to start imitating the voice of a person.

To conclude, based on the aforementioned research and the examples, there is a growing threat to many institutions, organisations, and other interests, as well as to society as a whole. What follows is an introduction to the main threats facing these groups separately, although it is important to state that each threat for one group is connected with the consequences facing another. From the view of government, deepfakes present multiple dangers. One is in military and national security: deepfake technology could generate false instructions and orders. From this perspective, deepfakes could be used as a form of disinformation supporting strategic, operational, and even tactical deception (Giles et al. 2019, 14; Chesney and Citron 2018b, 1783). At the same time, deepfake technology could support disinformation connected with the credibility of military services or intelligence agencies. Besides existing disinformation, which attempts to change real events (for example, Russia's information warfare), deepfake technologies could produce disinformation or authentic-looking news about an event which did not happen (or has yet not happened), for example, an attack against civilians in Iraq (Giles et al. 2019, 15; Chesney and Citron 2018b, 1783; Westerlund 2019, 40). All of these events could thus have consequences internationally—in international relations and diplomacy.

Moreover, deepfakes could also have an impact on democracy and trust in information spread by media and social media. For democracy, this manifests mainly in terms of elections (fake and manipulative political campaigns created by deepfake technology), the credibility of politicians (as was illustrated in the case of Obama), or the credibility of institutions

(attacks against judges, policemen, etc.). Another threat to democracy is connected with credibility and trust in media. Deepfake technology causes challenges not only for citizens in recognising what is real and what is fake, but also for journalists (for example, which eyewitness videos are real) (Chesney and Citron 2018b, 1779; CNN 2019, n.d.; Westerlund 2019, 42). The distrust could be even more dangerous than the deepfake itself. This distrust in media, news, and information can be called an 'information apocalypse' or 'reality apathy' (Westerlund 2019, 43).

For companies, the main threats are connected with credibility and privacy, mainly as it concerns transparency and data privacy. There are also dangers that some deepfakes will initiate fatal decisions based on totally false information (fake news, fake voice mails, etc.) which seems to be real. The threat of credibility is mainly connected with reputation (of companies or their leaders), which could be very easily damaged by deepfake technology (for example, by some fake impropriety videos). Another threat is connected with the manipulation of the market or manipulation of the decision-making process. Deepfake technologies could produce records which could be used for blackmailing or which would cause panic on the market (for example, news about a bankruptcy, etc.) (Westerlund 2019, 43).

For individuals, the first possible implication is anyone could be the target of manipulation, abuse, blackmailing, and so forth; this risk is higher for people who are celebrities, politicians, and the like. The second implication is that all content produced by us could be used as input for the creation of a deepfake (Giles et al. 2019, 17). This point is then connected with privacy and data protection. In connection with individuals, we also could point out that deepfake technology could be easily used in the abuse of children by child predators (Westerlund 2019, 43) or cause a public panic.

It is impossible to mention all the possible threats deepfake technology poses as the list is unlimited, but it is a real threat to the entirety of society. It is obvious that everyone is potentially a target of deepfake products, and, based on this, there are many challenges for researchers, politicians, and other experts to face. One of the challenges will be in the legal context of data privacy, data protection, and so forth. Another will be for IT experts to find tools for the detection and recognition of deepfake content (see Fraga-Lamas and Fernández-Caramés 2019; Nelson et al. 2020). And, last but not least, is a challenge not only connected with

deepfake recognition but also with other kinds of information manipulation: increasing the public's media literacy (Parkin 2019). However, it is important to note, that these threats stemming from deepfakes are not really novel. All the negative impacts on various institutions and parts of the society already exist and are posed by 'old' forms of faked content: primarily text and pictures. Deepfakes are simply expanding this to videos as well by making it so much easier to produce them. Just as it was possible to create doctored images long before using Photoshop and similar software, it was possible to create doctored videos before the advent of deepfakes. However, these technologies do make it massively easier, which leads to greater proliferation and production of such content. Whereas in previous decades manipulated videos were quite rare, in the coming decades they will probably become commonplace. The volume itself will become a problem. This prevalence will probably further erode public trust in information, making it easier for anyone seeking to destabilise and undermine society through the use of propaganda and information warfare.

2.7 Summary

The evolution of manipulative propaganda techniques goes hand in hand with technological development. With propaganda's 'going online', it uses voluntary as well as paid cyber troops, backed by the state and business, and primitive as well as sophisticated tools. Online propaganda manipulates people's behaviour in both nondemocratic and democratic countries. In nondemocratic regimes, the aim is to influence citizens and strengthen the regime. However, such countries can also have offensive aspirations against other state actors. As Russia's recent activities have shown, online propaganda and manipulation can be an important tool in power confrontation.

The current trends of modern online disinformation and propaganda activities were presented in this chapter from different perspectives. First was a focus on the information environment in the context of propaganda usage. New trends in nondemocracies, widened through the lens of the Russian example, were introduced. In Russia, the restriction of the Internet is still evolving. As mentioned, the newest example is RuNet, an isolated environment as of 2020 which manifests the 'digital sovereignty' of Russia. However, it does not mean that Russia or other 'isolated' countries will not influence and use propaganda tools upon others. Recent

manipulative Russian activities in Europe and the United States have demonstrated how sinister a virtual offensive can be.

The second part of the chapter was dedicated to the newest technological tools of online disinformation and propaganda and their fast evolution. Examples of troll and bot behaviour demonstrated how these originally primitive actors were able to evolve and become an important part in complex influence operations. Not long after the revelations of Russian influence activities in Europe and the United States, a new strategy in propaganda and manipulation appeared. The affairs surrounding the company Cambridge Analytica presented what influence the combination of big data, deep learning, and micro-targeting can have on voter behaviour. Moreover, Cambridge Analytica was hired primarily by actors in democracies, not undemocratic regimes, and not as a result of a confrontation between state powers.

In the last part, a new, future tool which may influence coming propaganda campaigns was introduced. Deepfake technology is based on artificial intelligence, and it is very likely that proliferation of this technology could be extremely serious as it relates to the credibility of all audio and video recordings (Schellmann 2017). The most concerning is its wide availability combined with speed and ease of production, leading to a trust-eroding deluge of deepfake videos.

In summary, the changing online environment and new tools using artificial intelligence (MADCOMs), deep learning, big data, and micro-targeting might be the future of propaganda and disinformation dissemination. However, given the quickly changing environment and rapid technological development, it is frankly not easy to predict what comes next. At present, it might seem that disinformation tools and propaganda capabilities are developing faster than countermeasures (Schellmann 2017). There is need for better threat recognition on the state level (see Chapter 7) as well as legal and strategic adjustments. Better technical abilities and understanding of defending actors are also needed. Media as well as technology organisations have a huge part to play in combating propaganda as well. They need to cooperate with state actors and realise that they also have a civil responsibility even if they work for business. Finally, there is an important role for society itself and particularly academia (see Chapter 8). Academics should still educate themselves about new technological developments and their possible consequences on propaganda even if it might be uneasy for their field of study. Their findings should be heard and discussed, especially by authorities who have the power to take decisive steps.

Bibliography

Al-Rawi, A., & Jiwani, Y. (2019). Russian Twitter Trolls Stoke Anti-Immigrant Lies Ahead of Canadian Election. *Public Radio International*. https://www.pri.org/stories/2019-07-26/russian-twitter-trolls-stoke-anti-immigrant-lies-ahead-canadian-election. Accessed 20 Mar 2020.

Aro, J. (2016). The Cyberspace War: Propaganda and Trolling as Warfare Tools. *European View, 15*(1), 121–132. https://doi.org/10.1007/s12290-016-0395-5.

Atik, H. (1999). *The Characteristics of the Information Economy*. Erciyes University Faculty of Economics and Administrative Sciences, Department of Economics.

Badawy, A., Addawood, A., Lerman, K., & Ferrera, E. (2019). Characterizing the 2016 Russian IRA Influence campaign. *Social Network Analysis and Mining, 31*(9). https://doi.org/10.1007/s13278-019-0578-6.

Bastos, M. T., & Mercea, D. (2018). The Public Accountability of Social Platforms: Lessons from a Study on Bots and Trolls in the Brexit Campaign. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376*. https://doi.org/10.1098/rsta.2018.0003.

Bernays, E. (1928). *Propaganda*. New York: H. Liveright.

Bowles, A. (2006). The Changing Face of the RuNet. In H. Schmidt, K. Teubener, & N. Konradova (Eds.), *Control + Shift: Public and Private Usages of the Russian Internet* (pp. 21–33). Norderstedt: Books on Demand.

Brikše, I. (2006). *The Information Environment: Theoretical Approaches and Explanations*. https://www.szf.lu.lv/fileadmin/user_upload/szf_faili/Petnieciba/sppi/mediji/inta-brikse_anglu.pdf. Accessed 26 Oct 2020.

Cadwalladr, C., & Graham-Harrison, E. (2018, March 17). Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach. *The Guardian*. https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election?CMP=share_btn_tw. Accessed 20 Mar 2020.

Cadwalladr, C., & Townsend, M. (2018, March 24). Revealed: The Ties That Bound Vote Leave's Data Firm to Controversial Cambridge Analytica. *The Guardian*. https://www.theguardian.com/uk-news/2018/mar/24/aggregateiq-data-firm-link-raises-leave-group-questions. Accessed 20 Mar 2020.

Capurro, R., & Hjørland, B. (2005). The Concept of Information. *Annual Review of Information Science and Technology, 37*(1), 343–411. https://doi.org/10.1002/aris.1440370109.

Chen, A. (2015, June 7). The Agency. *The New York Times*. https://www.nytimes.com/2015/06/07/magazine/the-agency.html. Accessed 20 Mar 2020.

Cheng, C., Kyungmin, K., & Ji-Yong, L. (2010). North Korea's Internet Strategy and Its Political Implications. *The Pacific Review, 23*(5), 649–670. https://doi.org/10.1080/09512748.2010.522249.

Chesney, R., & Citron, D. K. (2018a). Deepfakes and the New Disinformation War: The Coming Age of Post-Truth Geopolitics. *Foreign Affairs, 97*(6), 147–155.

Chesney, R., & Citron, D. K. (2018b). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *LAWFARE blog*.

Chessen, M. (2017). *The MADCOM Future: How Artificial Intelligence Will Enhance Computational Propaganda, Reprogram Human Culture and Threaten Democracy and What Can Be Done About It*. Washington: Atlantic Council.

CNN. (2019). Deepfake Videos: Inside the Pentagon's Race Against Deepfake Videos. *Cable News Network*. https://edition.cnn.com/interactive/2019/01/business/pentagons-race-against-deepfakes/. Accessed 20 Mar 2020.

Dickson, B. (2018, July). When Ai Blurs: The Line Between Reality and Fiction. *PC Magazine*, pp. 114–125.

Ebner, J. (2018, March 6). Forscherin schleust sich bei Hasskommentatoren ein - und erlebt Erschreckendes. *Online Focus*. https://www.focus.de/politik/experten/gastbeitrag-von-julia-ebner-hass-auf-knopfdruck-wenn-die-verbreitung-von-hass-computerspiel-charakter-bekommt_id_8554382.html. Accessed 19 Mar 2020.

Farrell, H. (2016, May 19). The Chinese Government Fakes Nearly 450 Million Social Media Comments a Year: This Is Why. *The Washington Post*. 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/. Accessed 19 Mar 2020.

Field, M., & Wright, M. (2018, October 17). Russian Trolls Sent Thousands of Pro-Leave Messages on Day of Brexit Referendum, Twitter Data Reveals. *The Telegraph*. https://www.telegraph.co.uk/technology/2018/10/17/russian-iranian-twitter-trolls-sent-10-million-tweets-fake-news/. Accessed 19 Mar 2020.

Financial Times. (2019, September 30). US Treasury Hits 'Putin's Chef' with New Sanctions.

Fraga-Lamas, P., & Fernández-Caramés, T. M. (2019). *Fake News, Disinformation, and Deepfakes: Leveraging Distributed Ledger Technologies and Blockchain to Combat Digital Deception and Counterfeit Reality* (Working Paper). Cornell University. https://arxiv.org/abs/1904.05386.

Freedom House. (2014). *Freedom on the Net: Russia*. https://freedomhouse.org/report/freedom-net/2014/russia. Accessed 20 Mar 2020.

Freedom House. (2018). *Freedom on the Net: China*. https://freedomhouse.org/report/freedom-net/2018/china. Accessed 20 Mar 2020.

Frenkel, S., & Wakabayashi, D. (2018, February 2). After Florida School Shooting, Russian 'Bot' Army Pounced. *New York Times*. https://www.nytimes.com/2018/02/19/technology/russian-bots-school-shooting.html. Accessed 20 Mar 2020.

Giles, K. (2016). *The Next Phase of Russian Information Warfare*. Riga: NATO Stratcom CoE.

Giles, K., Hartmann, K., & Mustaffa, M. (2019). *The Role of Deepfakes in Malign Influence Campaigns*. Riga: NATO Stratcom CoE. https://www.stratcomcoe.org/role-deepfakes-malign-influence-campaigns?fbclid=IwAR25Hbld-nb7nOaVmoW8AAnMLkiPxh8HoRRpjluIBhNWckD0yNZ8AJA1Sxg. Accessed 19 Mar 2020.

Godin, B. (2008). The Information Economy: The History of a Concept Through Its Measurement, 1949–2005. *History and Technology, 24*(3), 255–287. https://doi.org/10.1080/07341510801900334.

Gorwa, R., & Guilbeault, D. (2018). Unpacking the Social Media Bot: A Typology to Guide Research and Policy. *Policy & Internet*. https://doi.org/10.1002/poi3.184.

Grugel, J., & Bishop, M. L. (2013). *Democratization: A Critical Introduction*. London: Palgrave Macmillan.

Hannah, J. (2018). Trolling Ourselves to Death? Social Media and Post-Truth Politics! *European Journal of Communication, 33*(2), 214–226. https://doi.org/10.1177/0267323118760323.

Howard, P. N., Ganesh, B., Liotsiou, D., & Kelly, J. (2018). *The IRA and Political Polarization in the United States, 2012–2018*. University of Oxford: Computational Propaganda Research Project. https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/12/The-IRA-Social-Media-and-Political-Polarization.pdf. Accessed 19 Mar 2020.

Howard, P. N., Ganesh, B., & Liotsiou, D. (2019). *The IRA, Social Media and Political Polarization in the United States, 2012–2018*. University of Oxford: Computational Propaganda Research Project. https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/12/The-IRA-Social-Media-and-Political-Polarization.pdf. Accessed 19 Mar 2020.

Inkster, N. (2016). Information Warfare and the US Presidential Election. *Survival, 58*(5), 23–32. https://doi.org/10.1080/00396338.2016.1231527.

Jiang, M. (2010). Authoritarian Informationalism: "China's Approach to Internet Sovereignty". *SAIS Review, 30*(2), 71–89. https://doi.org/10.1353/sais.2010.0006.

Jiang, M., & King-wa, F. (2018). Chinese Social Media and Big Data: Big Data, Big Brother, Big Profit? *Policy and Internet, 10*(4), 372–392. https://doi.org/10.1002/poi3.187.

Kalathil, S., & Boas T. C. (2003). *Open Networks, Closed Regimes: The Impact of the Internet on Authoritarian Rule*. Carnegie Endowment for International Peace, Washington, DC.

Khondker, H. H. (2011). Role of the New Media in the Arab Spring. *Globalization, 8*(5), 675–679. https://doi.org/10.1080/14747731.2011.621287.

King, G., Pan, J., & Roberts, M. E. (2013). How Censorship in China Allows Government Criticism but Silences Collective Expression. *American Political Science Review, 107*(2), 1–18. http://j.mp/2nxNUhk. Accessed 20 Mar 2020.

Kollanyi, B., Howard, P. N., & Woolley, S. C. (2016). *Bots and Automation over Twitter During the U.S. Election: Data Memo*. University of Oxford: Computational Propaganda Research Project.

Kreiss, D. (2019). From Epistemic to Identity Crisis: Perspectives on the 2016 U.S. Presidential Election. *The International Journal of Press/Politics, 24*(3), 383–388. https://doi.org/10.1177/1940161219843256.

Libby, K. (2019, July 17). Giving Your FaceApp Selfie to Russians Is Really Bad Idea. *Popular Mechanics*. https://www.popularmechanics.com/technology/security/a28424868/faceapp-challenge-security-risks/. Accessed 20 Mar 2020.

Lorents, P., & Ottis, R. (2010). Cyberspace: Definition and Implications. In *Proceedings of the 5th International Conference on Information Warfare and Security* (pp. 267–270). Reading: Academic Publishing Limited.

Lu, J., & Zhao, Y. (2018). Implicit and Explicit Control: Modeling the Effect of Internet Censorship on Political Protest in China. *International Journal of Communication, 12*, 3294.

Maréchal, N. (2017). Networked Authoritarianism and the Geopolitics of Information: Understanding Russian Internet Policy. *Media and Communication, 5*(1), 29–41. http://dx.doi.org/10.17645/mac.v5i1.808.

Martens, B. (2015). An Illustrated Introduction to the Infosphere. *Library Trends, 63*(3), 317–361.

Mazarr, M. J., Bauer, R. M., Casey, A., Heintz, S., & Matthews, L. J. (2019). *The Emerging Risk of Virtual Societal Warfare: Social Manipulation in a Changing Information Environment* (p. 2019). Santa Monica, CA: RAND Corporation.

Nelson, S. D., Simek, J. W., & Maschke, M. (2020). Detecting Deepfakes. *Law Practice: The Business of Practicing Law, 46*(1), 42–47.

Neudert, L.-M. N. (2017). *Computational Propaganda in Germany: A Cautionary Tale* (Working Paper No. 2017.7). University of Oxford: Computational Propaganda Research Project. http://blogs.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/06/Comprop-Germany.pdf. Accessed 7 November 2019.

Nimmo, B. (2019). *Measuring Traffic Manipulation on Twitter*. University of Oxford: Computational Propaganda Research Project. https://comprop.oii.ox.ac.uk/research/working-papers/twitter-traffic-manipulation/. Accessed 7 Nov 2019.

Nycyk, M. (2017). *Trolls and Trolling: An Exploration of Those That Live Under the Internet Bridge*. Australia: Brisbane.

OED Online. (2020, October 21). "information, n.". Oxford University Press. https://www.oed.com/viewdictionaryentry/Entry/95568.

Parkin, S. (2019, June 22). The Rise of the Deepfake and the Threat to Democracy. *The Guardian*. https://www.theguardian.com/technology/ng-interactive/2019/jun/22/the-rise-of-the-deepfake-and-the-threat-to-democracy. Accessed 7 Nov 2019.

Pavlíková, M., & Mareš, M. (2018). Techniky robotické propagandy na sociální síti Twitter [Techniques of Robotic Propaganda on Twitter Network]. *Právo a Technologie, 9*(18), 3–28. https://doi.org/10.5817/RPT2018-2-1.

Porche, I. R., Paul, Ch., York, M., Serena, Ch. C., Sollinger, J. M., Axelband, E., et al. (2013). The Information Environment and Information Warfare. In *Redefining Information Warfare Boundaries for an Army in a Wireless World* (pp. 11–18). RAND Corporation.

Qi, A., Guosong, S., & Zheng, W. (2018). Assessing China's Cybersecurity Law. *Computer Law & Security Review, 34*(6), 1342–1354. https://doi.org/10.1016/j.clsr.2018.08.007.

Ristolainen, M. (2017). Should 'RuNet 2020' Be Taken Seriously? Contradictory Views About Cyber Security Between Russia and the West. *Journal of Information Warfare, 16*(4), 113–131. https://www.jstor.org/stable/26504121.

Robinson, O. (2018). *Malicious Use of Social Media: Case Studies from BBC Monitoring*. Riga: NATO Stratcom CoE. https://www.stratcomcoe.org/malicious-use-social-media-case-studies-bbc-monitoring. Accessed 7 Nov 2019.

Rose, H. (2019, May 29). Jessikka Aro, the Journalist Who Took on Russian Trolls. *The Sunday Times*. https://www.thetimes.co.uk/article/jessikka-aro-the-journalist-who-took-on-russian-trolls-fv0z5zgsg. Accessed 7 Nov 2019.

Schellmann, H. (2017, December 5). The Dangerous New Technology That Will Make Us Question Our Basic Idea of Reality. *Quartz*. https://qz.com/1145657/the-dangerous-new-technology-that-will-make-us-question-our-basic-idea-of-reality/. Accessed 9 Nov 2019.

Secureworks. (2017). Cyber Threat Basics, Types of Threats, Intelligence & Best Practices. https://www.secureworks.com/blog/cyber-threat-basics. Accessed 18 Feb 2020.

Segev, H., Doron, E., & Orion, A. (2019). *My Way or the Huawei? The United States-China Race for 5G Dominance*. The Institute for National Security Studies, Tel Aviv University. INSS Insight, no. 1193. https://www.inss.org.il/publication/my-way-or-the-huawei-the-united-states-china-race-for-5g-dominance/?offset=2&posts=2219&fbclid=IwAR1VgdtXd2msdotNHsOaCm_lkvyz_XFp4nDM0w0gjapp1DMIPdcVfQamTwE. Accessed 10 Jan 2020.

Solomon, S. (2018, March 22). Cambridge Analytica Played Roles in Multiple African Elections. *Voa*. https://www.voanews.com/africa/cambridge-analytica-played-roles-multiple-african-elections. Accessed 8 Nov 2019.

Spruds, A., Rožulkalne, A., & Sedlenieks, K. (2016). *Internet Trolling as a Hybrid Warfare Tool: The Case of Latvia*. Riga: NATO CoE.

Staudenmaier, R. (2018, October 18). Court in Finland Finds Pro-Kremlin Trolls Guilty of Harassing Journalist. *Deutsche Welle*. https://www.dw.com/en/court-in-finland-finds-pro-kremlin-trolls-guilty-of-harassing-journalist/a-45944893-0. Accessed 8 Nov 2019.

Steen-Thornhammar, H. (2012). Combating Censorship Should Be a Foreign Policy Goal. In J. Perry & S. S. Costigan (Eds.), *Cyberspaces and Global Affairs*. Burlington, VT: Routledge.

StopFake. (2019, June 9). Figures of the Week. https://www.shorturl.at/eNT04. Accessed 9 Nov 2019.

Tang, M., & Huhe, N. (2020). Parsing the Effect of the Internet on Regime Support in China. *Government and Opposition, 55*(20), 130–146. https://doi.org/10.1017/gov.2017.39.

The Computational Propaganda Project. (2016). *Resource for Understanding Political Bots*. http://comprop.oii.ox.ac.uk/research/public-scholarship/resource-for-understanding-political-bots/. Accessed 10 Nov 2019.

The Moscow Times. (2019, May 1). *Putin Signs Internet Isolation Bill into Law*. https://www.themoscowtimes.com/2019/05/01/putin-signs-internet-isolation-bill-into-law-a65461. Accessed 7 Nov 2019.

The State Duma. (2012). Legislative Measure No. 89417-6 2012, On Children's Protection Against Malicious Internet Content.

The State Duma. (2014a). Legislative Measure No. 428884-6, The Bloggers Law.

The State Duma. (2014b). Legislative Measure No. 553424-6, About Data Storage and Telecommunication Providers.

Thim, M. (2017). *Čínský internet pod rostoucí kontrolou* [The Chinese Internet Under Increasing Control]. Praha: AMO.

Titcomb, J. (2018, August 22). Facebook and Twitter Delete Hundreds of Fake Accounts Linked to Iran and Russia. *The Telegraph*. https://www.telegraph.co.uk/technology/2018/08/22/facebook-twitter-delete-hundreds-fake-accounts-linked-iran-russia/. Accessed 10 Nov 2019.

Toffler, A., & Toffler, H. (1981). *The Third Wave*. New York: Bantam Books.

U.S. Department of Justice. (2019). *Report on the Investigation into Russian Interference in the 2016 Presidential Election*. https://www.justice.gov/storage/report.pdf. Accessed 8 Nov 2019.

U.S. Joint Chiefs of Staff. (2006, December). *The National Military Strategy for Cyberspace Operations*. Washington, DC.

Wanless, A., & Berk, M. (2017). Participatory Propaganda: The Engagement of Audiences in the Spread of Persuasive Communications. In *Proceedings of the Social Media and Social Order, Culture Conflict 2.0 Conference*.

Wardle, C., & Derakhshan, H. (2017). *Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making*. Council of Europe. https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c. Accessed 6 Nov 2019.

Webster, F. (2006). *Theories of the Information Society*. Cambridge: Routledge.

Westerlund, M. (2019). The Emergence of Deepfake Technology: A Review. *Technology Innovation Management Review, 9*(11), 39–52.

What Is It. (2019). *Deepfake (Deep Fake AI)*. https://whatis.techtarget.com/definition/deepfake. Accessed 6 Nov 2019.

Whitehead, L. (2002). *Democratization: Theory and Experience*. Oxford: Oxford University Press.

Wolfsfeld, G., Elad, S., & Sheafer, T. (2013). Social Media and the Arab Spring: Politics Comes First. *The International Journal of Press/Politics, 18*(2), 115–137. https://doi.org/10.1177/1940161212471716.

Wondracz, A. (2019, August 6). Experts Warn to Stay Away from Data-Hoarding FaceApp—as 'Self-Absorbed' Selfies Could Lead to Devastating Privacy Breaches. *Daily Mail*. https://www.dailymail.co.uk/news/article-7324955/FaceApp-warning-experts-raise-concerns-privacy.html. Accessed 6 Nov 2019.

Woolley, S. C., & Guilbeault, D. (2017). *Computational Propaganda in the United States of America: Manufacturing Consensus Online* (Working Paper No. 2017.5). University of Oxford: Computational Propaganda Research Project.

Woolley, S. C., & Howard, P. N. (2016). Automation, Algorithms, and Politics| Political Communication, Computational Propaganda, and Autonomous Agents—Introduction. *International Journal of Communication, [S.l.], 10*, 9. ISSN 1932-8036.

Woolley, S. C., & Howard, P. N. (2018). *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media*. Oxford, UK: Oxford University Press.

Yonder. (2018). *The Disinformation Report*. https://www.yonder.co/articles/the-disinformation-report/. Accessed 6 Nov 2019.

Yuzeun, L., & Lyu, S. (2018). *Exposing DeepFake Videos by Detecting Face Warping Artifacts*. Computer Vision Foundation. http://openaccess.thecvf.com/content_CVPRW_2019/papers/Media%20Forensics/Li_Exposing_DeepFake_Videos_By_Detecting_Face_Warping_Artifacts_CVPRW_2019_paper.p. Accessed 6 Nov 2019.

Zhang, Y. C. (2017 [2016]) The Information Economy. In J. Johnson, A. Nowak, P. Ormerod, B. Rosewell, Y. C. Zhang. Non-Equilibrium Social Science and Policy. Understanding Complex Systems. Springer, Cham. https://doi.org/10.1007/978-3-319-42424-8_10.

Zhang, J., Carpenter, D., & Ko, M. (2013). Online Astroturfing: A Theoretical Perspective: Completed Research Paper. In *Conference: 19th Americas Conference on Information Systems*, *AMCIS 2013—Hyperconnected World: Anything, Anywhere, Anytime*.

Zhdanova, M., & Orlova, D. (2019). Ukraine: External Threats and Internal Challenged. In S. Woolley & P. N. Howard (Eds.), *Computational Propaganda: Political Parties, Politicians, Political Manipulation and Social Media*. Oxford: Oxford University Press.