FFI-RAPPORT

19/01194

Social network centric warfare

– understanding influence operations in social media

Arild Bergh

Social network centric warfare – understanding influence operations in social media

Arild Bergh

Norwegian Defence Research Establishment (FFI) 4 October 2019

Keywords
Sosiale medier
Informasjonsoperasjoner
Hybridkrigføring
Cyberforsvar
Russland

FFI report
19/01194

Project number
1368

Electronic ISBN
978-82-464-3229-8

Approvers
Ole-Erik Hedenstad, *Research Manager*
Jan Erik Voldhaug, *Research Director*
The document is electronically approved and therefore has no handwritten signature.

Summary

Since the early 2010s there has been a steady increase in the use of social media by states or terror organisations for the attempted manipulation of opinions and actions among people that are outside their legal control. Examples of this that have been reported in the media include the Russian hybrid warfare in the Ukraine and foreign interference in the US presidential election in 2016. For Norway, as well as other NATO countries, these developments represent a problematic new trend that requires different tools and skills to handle than what one has traditionally associated with influence operations. Although there is a large amount of documentation on these and other social media-based influence operations, little, if anything, has been done to try to explore *how* such campaigns might have an effect.

The aim of this report is to support the Norwegian Armed Forces and the Ministry of Defence to develop a better understanding of issues around social media-based influence operations. This is done by going beyond a mere summary of influence activities that have taken place in social media. Instead this report takes a socio-technical approach to examine various aspects of social media-based influence and contextualise them within studies of online social behaviours and general sociology and ICT-related research. For this purpose, the report uses secondary data from several cases of social media manipulation, both state-organised and smaller, more organic attacks.

From this base the report develops a conceptual chain that enables us to understand how an influence operation uses native aspects of social media to achieve its goals. In short, a planned influence operation is executed by active operators and relies on social media affordances (characteristics that facilitate certain activities). These affordances aid influence operations' amplification and reach so that the content is spread widely and is added to the continuously aggregated and accumulated content stored by social media services. This vast content collection is referred to as the *online information sediments*. This metaphor is used to emphasise the long-term, cumulative approach of social media where information never disappears but will fade in and out of view depending on what a user is interested in, what they are searching for, and so on. New content is affected by the online information sediments as existing posts will provide material for framing and understanding any new information. Alternatively, new posts may affect existing content by providing new ways of interpreting old posts. Either way, the information from influence operations competes for individuals' and groups' attention in an attempt to enter into and manipulate their meaning making processes. The aim would be to get targeted social media users to do something that is beneficial to the actor behind the influence operation.

Based on these new insights, the relevant authorities can start developing new approaches and procedures to detect, assess and possibly counter social media-based influence operations.

Sammendrag

Siden tidlig på 2010-tallet har det vært en jevn økning i staters eller organisasjoners bruk av sosiale medier for å forsøke å manipulere meninger og handlinger blant mennesker som er utenfor deres juridiske kontroll. Eksempler på dette som har blitt rapportert i mediene er den russiske hybridkrigen i Ukraina og utenlandsk innblanding i det amerikanske presidentvalget i 2016. For Norge og andre Nato-land representerer denne utviklingen en problematisk ny trend som krever andre verktøy og ferdigheter enn det som tradisjonelt har blitt brukt i forbindelse med påvirkningsoperasjoner.

Selv om det etter hvert er mye dokumentasjon på disse og andre påvirkningsoperasjoner i sosiale medier er det gjort lite for å utforske *hvordan* slike kampanjer kan ha en effekt.

Formålet med denne rapporten er å støtte Forsvaret og Forsvarsdepartementet i å utvikle en bedre forståelse av problemstillinger rundt påvirkningsoperasjoner i sosiale medier. Dette gjøres ved å gå utover en ren oppsummering av de påvirkningsaktivitetene som har funnet sted i sosiale medier. I stedet bruker rapporten en sosio-teknisk tilnærming for å undersøke ulike aspekter av sosiale medier-basert påvirkningsoperasjoner. Disse aspektene diskuteres så i sammenheng med mer generelle studier av sosial atferd på nett og sosiologi og IKT-relatert forskning. For dette formålet bruker rapporten sekundærdata fra flere tilfeller av sosiale medier-manipulering, både statlige og mindre, mer organiske angrep.

Fra denne basen utvikler rapporten en konseptuell modell som gjør oss i stand til å forstå hvordan en påvirkningsoperasjon bruker sosiale mediers iboende egenskaper for å oppnå sine mål. Kort oppsummert blir en planlagt påvirkningsoperasjon utført av aktive operatører og benytter egenskaper ved sosiale medier som legger til rette for visse aktiviteter. Disse egenskapene forsterker og øker rekkevidden til påvirkningsoperasjoner, slik at innholdet spres vidt og legges til i det kontinuerlig aggregerte og akkumulerte innholdet som lagres. Den enorme samlingen av nettinnhold blir kalt *online informasjonssedimenter.* Denne metaforen blir brukt for å understreke den langsiktige og kumulative tilnærmingen, hvor informasjonen aldri forsvinner, men er mer eller mindre synlig avhengig av hva brukeren er interessert i og søker etter. Nytt innhold påvirkes av online informasjonssedimenter fordi eksisterende innlegg danner materialet som påvirker tolkningen og forståelsen av ny informasjon. Alternativt kan nye innlegg påvirke eksisterende innhold ved å tilby nye måter å tolke gamle innlegg på. I begge tilfeller konkurrerer informasjonen fra påvirkningsoperasjoner om oppmerksomheten i et forsøk på å inngå i, og manipulere, meningsdannende prosesser hos individer og grupper. Formålet vil være å få brukerne til å gjøre noe som er gunstig for aktøren bak påvirkningsoperasjonen.

Basert på disse nye innsiktene kan relevante myndigheter begynne å utvikle nye tilnærminger og prosedyrer for å oppdage, vurdere og muligens motvirke slike operasjoner.

Contents

1 Introduction

This report consists of three chapters. The current chapter provides background information to social media in general and the use of social media to influence people. It also defines terms and concepts used in this report. This overview is followed in chapter two by an examination of current research and reporting on influence and influence operations in social media. This part develops a coherent conceptual framework to examine how influence operations use social and technical resources to spread their message and how such messages may achieve an effect by interfering in individuals' and groups' meaning making processes. The final chapter will summarise the ideas of this report and suggest areas that would benefit from further research.

1.1 Rationale for report

The Norwegian Intelligence Service's assessment of current security challenges mentions influence activities from Russia and other actors as an area of concern (Etterretningstjenesten, 2019, p. 12). The assessment also mentions new IT developments that provide new opportunities for such influence activities, in particular the "growth of non-editorial media platforms that systematically select news […]. It is challenging to counter these [platforms]" (Etterretningstjenesten, 2019, p. 15, author's translation). A survey of reports from eleven western countries' secret services suggests that Russian interference is using a divide and rule approach, and that social media is one of many avenues for their influence activities, albeit an increasingly important one (Karlsen, 2019). The ubiquity of the Internet in general, and social media in particular, across geographical and demographic boundaries give actors in a conflict the opportunity to try to affect opinions worldwide – at home; among the enemy, and perhaps most importantly; in third party countries, on a scale not seen before. As such influence attempts use a global, freely available infrastructure belonging to commercial companies it can be done at a comparatively low cost in terms of time and money. At the same time more and more people, particularly younger people, get their news first and foremost from social media (Stelter, 2008). Social media is therefore likely to be a key arena for possible influence operations in the foreseeable future.

From 2014 to 2016 the use of social media as a component of state and terrorist level conflicts can be said to have matured in use and simultaneously entered the public consciousness (Goolsby, 2019). From the Middle East the ISIS organisation successfully used videos on YouTube and other social media to spread propaganda and recruit fighters from around the world to their cause (Bardin, 2014; 'Cyber Caliphate', 2015; A. Fisher, 2015; Matejic, 2016; Mazzetti & Gordon, 2015). Russia annexed Crimea from Ukraine in 2014 and used social media to present narratives such as Crimea being helped by Russia in a crisis or Crimea belonging to Russia (DeRosa, 2015; Dougherty, 2014; Geers, 2015; Treverton & Miles, 2014). And in the US presidential election of 2016 Russian operatives actively engaged in influence operations that sought to exploit existing divisions in US society to favour certain candidates they believed would be more positive for Russia (McKew, 2018; Parlapiano, 2018). It has been argued that for Russia such information operations are not necessarily an appendix to more traditional state

power tools, but that *"Russia will consistently use information operations as an independent, decisive tool of statecraft"* (Allen & Moore, 2018, p. 69). Such operations may have roots going back many decades in the Soviet state apparatus, and is now, as Abrams says *" […] through the Internet, able to influence popular opinion on a scale never before possible"* (Abrams, 2016, p. 8). Influence operations in social media is a field that will probably be fluid for a long time. Counter efforts by different social media actors (whether the company controlling it, or volunteers) has resulted in changes from the influence operation, which result in new counter measures (Fandos & Roose, 2018). Examining a single case to prepare for future influence operations will thus be of limited value. The aim of this report is therefore to develop new tools in the form of concepts that can help broaden our general understanding of this relatively new field for influence operations.

Given the conditions outlined above it is now problematic for states and armed forces to ignore what is happening in social media. The level of recent activity suggests that adversaries believe their efforts to be useful. Moreover, foreign influence operations may be a breach of political sovereignty regardless of how effective they truly are. In this sense, any state has a vested interest in reducing foreign influence. The rationale for this report is thus to start a process of understanding such operations on a deeper social and technical level. From this base one can discern future directions of research to pursue in order to develop practical skills and knowledge that is of use in this nascent field of social media based influence operations. The key contribution of this report is the suggestion that influence operations in social media can be conceptualised as a chain of tools and arenas that are connected through activities, with the central arena being individuals' and groups' meaning making processes.

1.2 Approaches, goals and limitations of this report

This report will argue that to prepare for possible adversarial influence operations in social media a socio-technical perspective will be beneficial. This perspective suggests that the social and technical aspects are not separate. Rather one should explore social interactions from a social sciences perspective and how those interactions are influenced and facilitated by technology, but also see how social aspects in turn affect technology and technology developments. As social media is a fusion of technology and social interactions the socio-technical approach will enable a deeper understanding of the processes that take place when someone seeks to influence social media users. There exists a fair amount of research focusing on summarising and reporting influence operations in social media (some examples from NATO include Geers, 2015; Bialy, 2017; Svetoka, 2016). However, there is so far not a lot of in-depth research connecting social media influence operations with the ways in which people interact with and process information and the technical underpinnings of social media.

This report undertakes a socio-technical analysis to address this issue by examining recent information on, and academic examination of, influence operations that have taken place in social media since the early 2010s. Findings in this literature are then explored in combination with sociological and technical literature on relevant topics, such as *attention online* or how *algorithms* (automated software routines) can be misused to move beyond reporting on what has happened and try to see how social media based influence operations may have some effect.

The focus of this report is on what happens in and around social media. The larger picture in which an influence operation takes place (political issues, other forms of influence operations, kinetic warfare, etc.) is outside the scope of the discussion here. Related online attacks, such as phishing for confidential emails or hacking the power grid, is also outside the scope of this report.

A key limitation of this report is that a lot of the data examined relates to other countries. However, if the same actors are involved on the attack side as seen in many of the cases used in this report, it is fair to assume that they will probably apply similar techniques, such as the pollution of the information space rather than clearly defined influence operations (Richter, 2017). One may therefore assume that many of the issues in existing cases will apply also in a future, Norwegian scenario. Furthermore, influence operation tactics that have been used abroad are quickly adopted and adapted locally. For instance, the deliberate creation of *fake news* stories to earn money on divisive issues is now done by a network of Norwegian websites (Bergsaker & Bakken, 2018).

Another limitation of this report relates to the newness and agility of the field. In terms of building up deeper understanding of what is going on these two aspects are limiting us considerably. This report is therefore doing an initial mapping of the terrain of social media and influence (operations) related issues as it is today. Finally, this report examines influence operations aimed at the general population, including the military. It does not look at efforts directed specifically at armed forces such as texting soldiers in Ukraine (Satter & Vlasov, 2017).

1.3 Examples used in this report

Reporting on social media influence operations commonly focuses on issues such as social media platforms used; what tools (such as bots) were applied; how many people saw posts from these operations or what content was presented. This article wants to move beyond such summaries to try to understand how social media influence operations may have an effect on people, and to do so from a user oriented perspective. However, obtaining data from online influence operations can be legally challenging and requires considerable time. This report is therefore using secondary data, primarily from three well-known social media influence operation cases, as an input to illustrate and illuminate issues under discussion. These issues are explored further through additional sociological and technical literature that expands our understanding of the processes in such influence operations.

1.3.1 Example 1: Annexation of Crimea

A key wakeup call for the military was the deployment of social media as a *force multiplier* (Giles, 2015, p. 4; Herrick, 2016, p. 111; Perry, 2015, p. 5) during the annexation of Crimea. The operation has been described as hybrid warfare (Hansen, 2017; Svetoka, 2016), where actual kinetic force was a smaller component of the attack. As Perry suggests *"[hybrid warfare's] successful use ultimately relies on an effective information operations campaign supplemented by coordinated special operations conducting unconventional warfare"* (2015, p. 2). Giles (2015) and Iasiello (2017) suggest that Russian information capabilities and social media use were deliberately honed in response to several key conflicts where the Russian state felt that they were on the losing side of the information war. These experiences *"led to the conclusion that automated systems are simply not sufficient, and dominating mass consciousness online requires the engagement of actual humans"* (Giles, 2015, p. 3). This is an important point to note, as it goes against longstanding beliefs (rooted in economic reasons) by social media operators such as Facebook or Twitter, that automated tools can alone mitigate information attacks (see e.g. Claburn, 2017; Kastrenakes, 2016).

1.3.2 Example 2: US Election 2016

The United States presidential election in 2016 is of particular interest due to the wide variety of *types* of actors involved; most using digital tools to exploit and amplify existing social divisions (Penzenstadler, Heath, & Guynn, 2018) for their own purposes.

The (initially unexpected) winner used social media more directly than previous and other candidates to bypass gatekeepers in other media channels. Supporters of different candidates, sometimes augmented by semi-automatic tools (Timberg, 2017), were also very active in social media, often spreading information later proven to be false.

Different commercial actors in three broad categories were also involved, albeit unwittingly or within the rules of regular, domestic political campaigning or for commercial, non-ideologically reasons. 1) the social media platforms such as Facebook, Twitter, 2) consultants for candidates who profiled and targeted social media users and 3) smaller, commercial actors interested in earning money, usually by pandering to extreme views by creating fake news to sell adverts (Kirby, 2016; Wendling, 2018).

In the aftermath it has become clear that actors linked to the Russian state utilised the openness of (western) social media to fuel the distrust and rage that emerged from the election season, using a so-called *troll army* (Hern, Duncan, & Bengtsson, 2017; Higgins, 2016; Seddon, 2014). This was done through creating fake news, spreading rumours, using bots to inflate viewer and share counts on different social media services, buying adverts to spread certain viewpoints, etc. (Penzenstadler et al., 2018; Oremus, 2016; Poulsen & Ackerman, 2018; Timberg & Romm, 2018; Shane & Mazzetti, 2018; Brandom, 2018; Timberg, Dwoskin, Entous, & Demirjian, 2017; Devine, 2017; O'Sullivan, 2017; Walker, 2015).

1.3.3 Example 3: Gamergate

Gamergate (or #gamergate) was a loosely organised attack on women in gaming using primarily social media. It emerged from aggressive responses to a game written by Zoe Quinn (Parkin, 2014b). Quinn, and other women who participated in online discussions relating to gamergate itself (Hern, 2014; Robertson, 2014; Valenti, 2015) or the issues gamergate ostensibly cared about, were threatened with (sexualized) violence. Attackers claimed they wanted to discuss ethics around gaming and media, in reality they focused on harassment of individuals (PM, 2014), the discussion being "a pretense to make further harassment of women in the industry permissible" (Parkin, 2014a).

Figure 1.1 Tweets from #gamergate.

1.4 Concept and terminology clarifications

This report will use the concept definitions below, these are specific to this report and may, or may not, mirror more broadly available definitions (for more traditional definitions cf. e.g. Adams, Brown, & Tario, 2009; Larson et al., 2009; Nicander, 2001; Santa Maria, 2013; Treverton, 2017).

1.4.1 Influence operations defined in the context of social media

An *influence operation* in social media is the attempt by an *initiating actor* to interfere in the process of meaning making among a *target audience* outside their legal control by generating and/or distributing information through openly available social media platforms. A *defending actor* may attempt to stop or reduce the impact of such operations, whereas individuals or groups that generate and/or distribute the original or related information, but are not directly controlled by either side, are *third party actors*. It should be noted that by designating a concerted effort as an influence operation, influence is not guaranteed and any influence that takes place may not be what was intended. It is the attempt at influence that is covered by this term.

An example of an influence operation would be the deliberate attempt by country A (the initiating actor) to encourage the belief among citizens in country B (the target audience) that a certain region of their country historically belongs to country A. The long term goal could be to take over control of the region, possibly through coercion or hybrid warfare. The authorities in country B (the defending actor) may attempt to counter this by, for example, posting counter-information. A commercial company (i.e. a third party actor) may create fake news stories supporting country A, not because they share the same ideology, but to earn money from advertising around the fake news.

Information used by the initiating actor may be truthful or falsified, but is selected so as to encourage viewpoints that support their ultimate goal, whatever that may be. Information may be created and posted directly, but it could also emerge by someone performing activities that result in information that support the initiating side being created by others. For instance, hacking and closing down a power station would cause news outlets and others to discuss that event online. The overall aim for the initiating side is to provide input for the generation of meanings among the target audience that are favourable to themselves and non-favourable to the other side. During such influence operations in social media all four types of actors actively engage with the information. This is unlike a traditional influence operation where the target audience would receive information but could not interact with it and third parties would be largely absent.

The use of the influence operation term in this report will generally relate to operations executed through social media. When discussing influence operation in general terms this will be clear from the context of the discussion. The term meaning making is from social sciences (Krauss, 2005) and highlights the fact that different actors' reaction to information is not pre-determined. The "Stab-in-the-Back Legend" that emerged after Germany's loss in World War I, and a similar response to the USA pulling out of Vietnam in 1975 (Kimball, 1988) exemplify this. These legends claim Germany/USA would have won but was stopped by people on their own side. This was a result of some people looking for acceptable explanations to a, for them, emotionally problematic development. Meaning making depends on *framing*. In short, *"framing theory suggests that how something is presented to the audience (called 'the frame') influences the choices people make about how to process that information. Frames […] work t o organize or structure message meaning."* (Davie, 2011; see also Bjørnstad, 2019 for a more detailed discussion).

1.4.2 Social media defined

At a glance it may seem that social media equals Facebook and Twitter plus the flavour of the month (Pinterest, TikTok, etc.). However, before deciding on what to include in this research, social media needs to be defined more clearly:

> Social media are services available through the Internet that allow the posting of content by people who do not operate or control the service; and the facility for other people to access, use and respond to such content.

This definition is broad, but it allows us to perceive online, social content related interactions in other places than just traditional social media. This includes comment fields in local newspapers and online reviews as well as Facebook and Flickr.

1.4.3 Other terms

Algorithm: An algorithm, in computer terms, is a process for solving a (often repeated) problem, following fixed steps. To detect spam email, for example, a number of different algorithms are applied. The terms have become known to the general public in relation to algorithms that make recommendations, such as selecting news they think you will like based on previously read stories or showing adverts based on your shopping cart at Amazon.com.

Artificial intelligence / Machine learning: Machine learning (or ML for short) is one of many forms of artificial intelligence. Machine learning refers to software that is developed in such a way that it learns without human supervision. In this report I will generally use machine learning because it reflects the main tasks when trying to detect influence operations in social media, i.e. train software to recognise patterns in text or images such as finding large numbers of negative messages about a particular politician. Artificial intelligence will only be used when discussing a topic where the term has been used by others originally.

Bot: Short for robot, this is automated software (i.e. not a human-like physical robot) that is used to interact with social media accounts. It may *like* all posts that mentions a particular topic or it can retweet it to other bots, and so on, thus creating the impression that the topic in the post is important.

Typology: Several of the discussions in the report will cover topics that have been reported extensively on in other media. However, these discussions often treat the core topic as a simple either/or proposition. For example fake news is only labelled as such, without further examination. In some cases this report will introduce **typologies**. Typologies is a useful analytical tool to organise elements (Doty & Glick, 1994; Smith, 2002) and get away from such binaries, and present the matter in a more nuanced way that will facilitate a more realistic discussion about the topic at hand. Real life is a messy affair; reducing complex phenomena to one-dimensional shorthand expressions will therefore not aid us in handling issues appropriately.

2 Social media influence operations: Understanding the people – technology nexus

To understand how influence operations can unfold on social media this report suggests that it can be conceptualised as a chain of tools and arenas that are connected through activities.[1] Activities can be done by the attacker; by automated, technical routines or by third parties who may or may not share in the goals of the attacker. Examples of third parties can be individuals earning money from fake news relating to the topic(s) of the influence operation (Wendling, 2018) or groups inside the country being attacked supporting the attacker (Higgins, 2016).

In table 2.1 below this conceptualisation is listed stage by stage. References in brackets point to the section in this chapter where the stage is discussed in depth. The last column shows examples of socio-technical aspects of social media that each stage relies on. A summarised version of this chain is found in chapter 3.

[1] Facebook's security department has discussed some of these stages in the context of Facebook having been used for information operations (Weedon, Nuland, & Stamos, 2017). This conceptual chain develops these ideas further to fully account for the social and technical interdependencies.

Conceptual chain	Relying on, for example:
A planned influence operation executed by active operators [⇨2.2]	- Content creation - Content dissemination
relies on Affordances of social media [⇨2.3]	- Anonymity - Geographical distribution
that aids the Amplification and reach [⇨2.4]	- Trolls - Selection algorithms - Searches - Reuse, promotion
which contributes to the Online information sediments [⇨2.5]	- Current & past information - Synthesised through algorithms, curation
that are deployed to fight for Individual or group attention [⇨2.6]	- Influencing - Derailing
to manipulate Individuals' or groups' meaning making processes [⇨2.7]	- Building or maintaining personal world view - Narratives
so as to encourage Alternate individual or group (in)actions [⇨2.8]	

Table 2.1 A conceptual chain of activities, events and arenas in a social media influence operation.

2.1 Overview of the conceptual chain

In the conceptual chain shown in table 2.1, an *active operator* (a nation state or terror group for instance) executes a *planned influence operation* by creating content for an initial intervention on social media. The material may be emails from a hacked account, a fake, topical video or even just a re-framing of news from an official source. Social media provides certain *affordances* to anonymously distribute the content and also facilitates further *amplification and reach* to subsidiary audiences such as followers of the profiles used to post the content or by traditional media picking up on the content. The sum of new social media posts and existing, related, posts, as well as additional posts that are triggered by the intervention together form *sediments of information*; a vast set of data that is publicly available through the Internet to stir up for different purposes.

These information sediments become resources for gaining *individuals' and groups' attention* through posts that social media users either seek out or have automatically added to their social media feeds. For example, if a social media user are already prone to believe that the Norwegian child protection services are "acting like Nazis" (Christopoulou, 2018), other posts that confirm this view are a) likely to show up in their searches or social media feeds and b) may be incorporated into their world view(s). Thus the influence operation's intermediate target is for these posts to *influence users' meaning making processes*, with the ultimate aim being to have

such posts present in convincing qualities and quantities that they may reach a tipping point and push the users into *alternate (in)actions*, i.e. actions they would otherwise not take. This may be to abstain from doing something, such as not voting for a party or not voicing an opinion or doing something active like protesting against, or voting for, certain measures.

2.2 A planned influence operation

The most common approach seen in the influence operation examples used in this report relies on exploiting contentious issues among the target population that can be used to push a certain agenda or sow divisions. Examining relevant issues and creating content about these issues are therefore key parts of a social media based influence operation. Hackers may also try to get hold of data that can be used to launch the campaign, for instance compromising material about politicians in the target area. Posts may also be trialled in a similar way to dark ads, ads that are only visible to the advertiser and an audience specially selected by profiling them (Sloane, 2016), and which, it has been suggested, can affect individuals political opinions (Hern, 2017b). Others will not see these posts and the attack trials are undetected, and it has been noted that *"the preparatory phase of hybrid warfare does not differ that much from the conventional tools of Russian diplomacy"* (Rácz, 2015, p. 73).

Apart from text, manipulated images and animations/videos are created with software before the content is pushed out through accounts on different social network outlets, either manually or through automated software and the popularity may be inflated through bots that *like* or *repost* the content. This combination of accounts and software plus the affordances of social media that will be discussed below provide a *soft infrastructure* that supports and facilitates the influence operation. An example of how it works as infrastructure can be seen in an operation where bots that had previously been used to post pro-Palestinian content were re-used for pro-Brexit content (Howard & Kollanyi, 2016).

It is important to be aware that although an influence operation may be social media based, it can move offline and have effects in the real world. An examples of this include 'Pizzagate', where a conspiracy theory spread online claimed paedophiles linked to Hillary Clinton abused children in a pizza restaurant in Washington D.C., and a man shot into the restaurant believing this to be true (Siddiqui & Svrluga, 2016; Editorial Board, 2016). Furthermore, Russian-linked group were found to sell *Black Lives Matter* merchandise online to further fan polarisation in US politics (O'Sullivan, 2017). Individuals under threat in online attacks are often victims of *doxxing*, this is where their real life location (home or work) is revealed online, implicitly to make physical threats to them more believable (Bowles, 2017; Klang, 2016; Molden, 2015; Sinders, 2015).

What differentiates a social media-based influence operation from other social media activities that try to influence users is that a) it is initiated and (in part) directed by a larger organisation or state actor; b) it has ultimate goals that it hides from the target population. What distinguishes a social media-based influence operation from other influence operations is that a) it avoids any need to use intermediaries; b) the content mixes with other, mundane content and c) the content

created is accessible outside the influence operation context afterwards. The two latter differentiators are also examples of what will be discussed next, namely affordances of social media.

2.3 Affordances of social media

The second element of the conceptual chain above (table 2.1), affordances of social media, focuses on how different types of social media will facilitate individual influence operations approaches and actions in distinct ways. The concept that objects afford certain actions based on how they are perceived was developed by Gibson (1977, p. 127). To illustrate the affordance concept one can use a chair. It affords sitting (more than other actions) by the way it is shaped and a shared, local understanding among actors of what a chair is. It may however, also be used as a projectile or a ladder. Thus the affordances of social media for example are not static, but represent the most likely action(s) of many that they can be used for (Bergh, 2015).

Using a basic example regarding social media one can see that a blog affords longer engagement, longer lifetime of content and a slower spread (if any) through available sharing tools. Twitter on the other hand would facilitate a rapid spread, less time to engage critically with content and a shorter time in which the content is displayed prominently in a user's news stream. The affordances of social media are thus an important element when it comes to deploying or countering influence operations in social media. In figure 2.1 we see screenshots of two very different types of social media that illustrate such affordances.

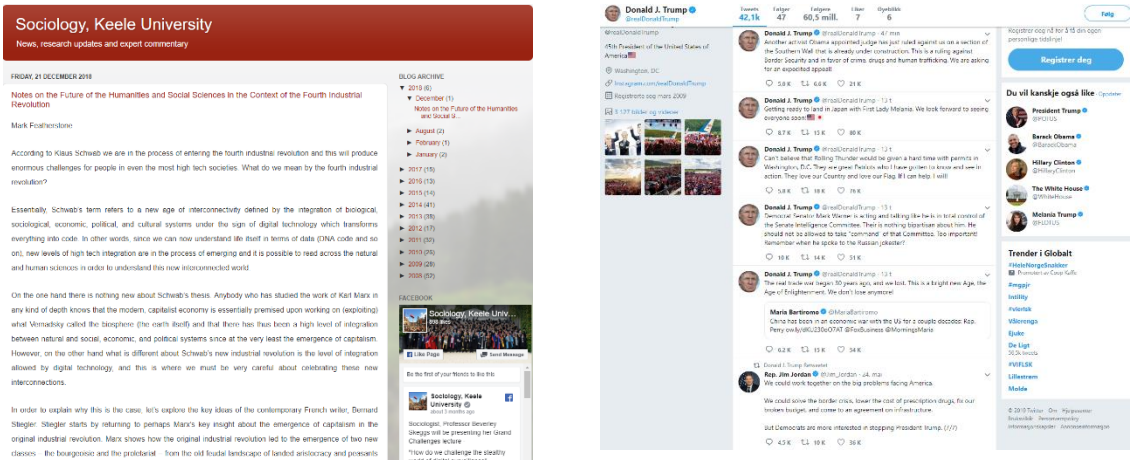


Figure 2.1 Left: Typical blog with long text and a list of previous entries on the right. Right: Twitter stream; short texts and interaction buttons, no historical navigation.

When examining affordances of social media in the context of influence operations there are many dimensions to consider. One dimension is the type of media (short text, video, image, mixed content, etc.) that the social media service uses. The rise of visual information on social media ties in with this. This is partly fuelled by constant access to smartphone cameras and partly through specialised social media that focuses on sharing videos or pictures (Walter,

2012). The use of visual information is frequently used to influence people, whether for political (Seo, 2014) or commercial gain (Kerwood, 2016). Pictures have habitually been used in online influence operations (Timberg & Romm, 2018; Permanent Select Committee on Intelligence, 2018), often taken out of context and/or altered; see figure 2.2 for an example of this.

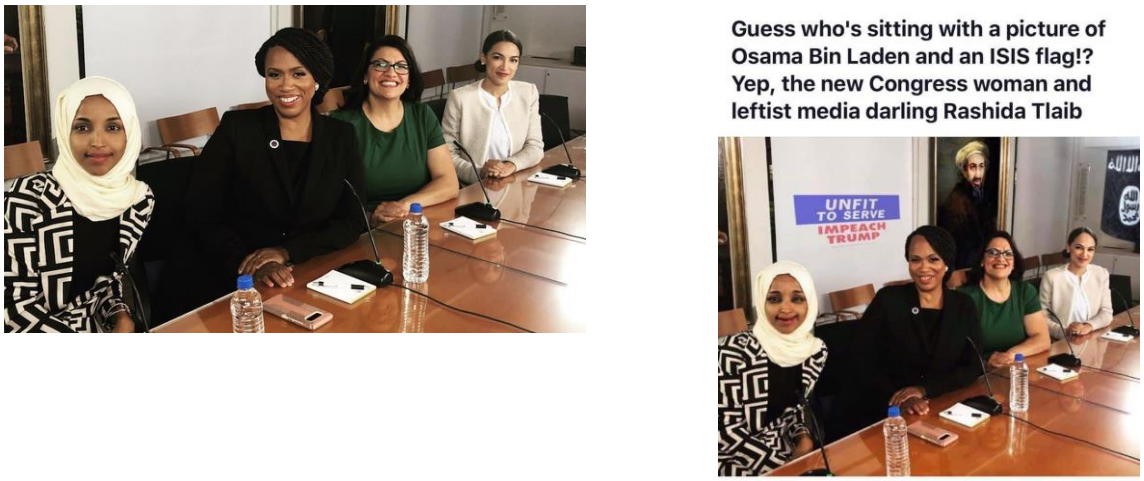


Figure 2.2 Left: original image showing four new congresswomen in the US (Folley, 2018). Right: Image doctored to suggest link to ISIS and Bin Laden (TKOBeauty, 2019).

Images can have a stronger appeal to emotions that bypass cognitive work (see for example Joffe, 2008; Richardson & Wodak, 2009; Seo, 2014). In an analysis of the use of Twitter by the Israeli Defence Force and Hamas during the November 2012 conflict, it was found that Hamas used much more emotional images, appealing more directly to the user (Jolicoeur & Seaboyer, 2018). These affordances can help an attacker to create posts that get more attention through appeals to emotions. Visual information is also vaguer, this can be very positive when trying to persuade people, as Richardson and Wodak expresses it when analysing the visual language of racist pamphlets "*we believe vagueness to be an inherent feature of political communication and also for advertising, particularly in images or metaphors*" (2009, p. 51). Such vagueness also allows social media users to project what they want, an example of this occurred on a Norwegian closed Facebook group against immigration where empty bus seats were taken to be Muslim women wearing burkas (Henriksen, 2017).

The emergence of so-called *deepfake* videos will only exacerbate this in the future; such videos are created by altering parts of a video, for instance replacing the head of someone in a sex-video with the head of a political leader. The quality of these fakes is such that it becomes difficult to see that it is not real. Time-based media, such as audio and video, also makes it more difficult in general for the defending side to detect an attack as considerable more computing resources are required to analyse the content of a video than a text message. Finally, visuals also bypass the need for language skills if an influence operation is geared toward speakers of a different language, it is a cheap and simple way to cross cultural barriers (Seo, 2014).

Another dimension of social media affordances relates to the capabilities of devices used to interact with social media, for instance PCs, smartphones or tablets; the potential audience one can reach is another. These dimensions all shape an influence operation's reception. An unemployed activist on a PC can follow links, skim articles and respond faster than a busy commuter checking news on the way home. The latter may find it easier to share content due to smartphone sharing facilities. Thus the ease of sharing content, as a capability of smartphones, directly affects how content spreads. In India for example, the WhatsApp messaging platform felt it necessary to restrict the ability to share message in response to false rumours that were spread through smartphones that directly led to attacks on individuals causing several deaths (Cellan-Jones, 2019). Thus if one plans to use copies of reputable news websites with fake news stories (Ruddick, 2017), targeting smartphone users with limited time that quickly glance at something in a feed and sees what seems to be a reputable source may have a better chance of spreading.

Social media as influence operation platforms

On the whole one can say that social media represents a *platform* for the influence operation. In information and communications technology, *"a platform is a group of technologies that are used as a base upon which other applications, processes or technologies are developed."* (Techopedia, 2019). Facebook is seen as a marketing platform, for instance. However, this report would argue that all social media services in combination are a single platform for influence operations. For example when YouTube banned videos showing how to make your own guns, some of the former YouTube users moved their content to a pornographic video website (D. Lee, 2018). And extremist such as ISIS who had their accounts closed on Facebook or Twitter moved to other, smaller sites without capabilities for moderation, such as Justpaste.it (Fishwick, 2014; Hern, 2017a). The implication here is that removing adversarial content from an influence operation on one service, if at all possible, is an ongoing process and not a checklist item that can ever be considered complete.

A real life example of how social media affordances affect information reception is the Ukrainian soldiers who received direct SMS messages threatening them in the field. It was suggested that this is a form of pinpoint propaganda and that *"[t]here's just something about viewing a message on your phone that just makes people more susceptible or vulnerable to its impact"* (Nancy Snow quoted in Satter & Vlasov, 2017). These messages were sent directly to their personal devices, in the same format that they would receive everyday messages from friends and acquaintances, sometimes even pretending to come from fellow soldiers. The importance of getting your message into a common social media stream was recognised by organisers of Barack Obama's 2012 US presidential campaign who claimed that

> *[u]nderstanding that a message from a friend is more trusted and effective, the program mmessages boosted target audience reach by 400% and increased completion rates for important actions like registering to vote by 40%.(Pilkington & Michel, 2012)*[2]

[2] The psychological explanations for this focus on issues of ingroup effects, concencus and persuasion; see *Understanding communication and influence in a defense context: A review of relevant research from the field of psychology* (Bjørnstad, 2019) for more on this.

Algorithms constitute another dimension of social media affordances and will be discussed in detail in section *2.4.1 Amplifying by gaming the technology*. At this point it is just important to highlight that the ease with which one can manipulate a social media service's algorithms to reach many users and/or create the impression of acceptance or popularity are affordances that an attacker would probably consider when starting an attack on a particular social media service.

In sum, affordances of social media represent resources for an influence operation which can affect how information is received, processed and shared and thus contribute to adversarial information entering individuals' and groups' meaning making processes.

2.3.1 Social media differentiators

Partly overlapping with affordances are the properties that make social media unique in comparison to traditional, offline mass media. These differentiators, such as worldwide reach, are generally well known, but it is worth briefly examining them here in the context of social media based influence operations.

- **Reach**: The instantaneous and non-geographical nature of the Internet is the key characteristic of social media with potentially immediate, worldwide dissemination of influence operations material. Lack of geographical boundaries means that actions intended for one audience can trigger or reach other audiences, causing collateral damage. A Russian politician may boost their family policy credentials locally by criticising the Norwegian Child Services (Mørch, 2014), a controversial subject in Eastern Europe. Even if intended only for a Russian audience it will reach others through social media and thus contribute to negative perceptions about Norway. Conversely, manual efforts by social media services to moderate content use centrally located staff (M. Fisher, 2018; Hopkins, 2017), posts from one location are consequently examined by people in a different place with dissimilar tacit knowledge and another cultural understanding than the intended audience, making moderation more difficult.

- **Use and control**: In liberal democratic societies there is no central, national oversight or control over social media, and social media services are often legal entities based in another country. This openness makes it easier to exploit social media platforms to conduct influence operations (e.g. establish soft infrastructure, spread disinformation). Furthermore, many Internet services remove intermediaries to lower costs and provide individuals or small groups with opportunities that have previously only been available for large or national entities, for instance global video distribution. These opportunities are widely used by fringe groups to push their agendas beyond their core groups. ISIS is a recent example of this, they used YouTube as well as smaller websites to spread their propaganda (cf. e.g. McHugh, 2015; Fishwick, 2014; Mazzetti & Gordon, 2015).

Figure 2.3 Left: Photocopied 1970s neo-nazi pamphlet (Worley & Copsey, 2016). Right: Current neo-nazi website.

- **Reception and impression**: Reach and access for information from an influence operation is further helped by the fact that one cannot necessarily distinguish *visually* between a website (or a social media news item) from a reputable news organisation, and someone attempting to spread fake news. This approach, and the reasons for this, is discussed in some detail elsewhere (Alme, 2019). There is little to no cost involved in creating a website that looks like a newspaper, it is even cheaper to simply copy the entire design of an existing website. This has been done and combined with using domain names that look similar to the original news sites domain name and then used to create fake news articles that seem legitimate when shared through social media (Ruddick, 2017). This is in stark contrast to pre-Internet communications. In figure 2.3 an example of a 1970s neo-Nazi pamphlet is compared with a modern neo-Nazi website. On social media news streams there is a single, uniform look to all items, hence the platform itself flattens content and removes cues that could help readers differentiate between news items, the way they would do in real life. Furthermore, the anonymity of users on the Internet makes it difficult for users to discern who is spreading fake news, whilst pseudonymity (Tsikerdekis, 2012) allows users to claim expertise or knowledge that they don't possess on certain topics, or by co-ordinating multiple online identities they can make supporting statements to enhance the believability of information presented.

- **Content accessibility – devices**: The rapid spread of smartphones and tablets is important here; audiences that before were difficult to reach, such as soldiers in the field, are reachable in real-time. Furthermore, the size of these devices in terms of reading content and their frequent use as secondary devices to consume content on the side of a main activity (Shin, An, & Kim, 2016; Van Cauwenberge, Schaap, & Van Roy, 2014) means that content may not be scrutinised much before acting on it (Gabielkov, Ramachandran, Chaintreau, & Legout, 2016).

- **Content accessibility – selection**: Most news and social media sites rely on algorithms to automatically select content they believe is relevant and/or desired by the user (Bucher, 2012; Dias, 2014). This is done to increase the stickiness of the site, that is, the amount of time spent engaging with the website. At the same time there are specialist forums and groups on any imaginable topic which results in a self-selection bias with regard to what content one accesses (Bakshy, Messing, & Adamic, 2015, p. 1130). One may also see *self-deselection*, where moderate users leave online discussions due to aggressive responses to their postings, a tactic that has been used in Russian influence operations (Sindelar, 2014).

- **Content accessibility – access to content over time**: This is a major difference from traditional media; one that this report suggests presents a major resource for attackers. On social media, unless specifically designed for it, nothing disappears, it merely fades from view. Facebook for instance uses cold storage to keep all photos ever uploaded, with each facility able to store in excess of 1 Exabyte, equalling approximately 10 trillion photos (Bandaru & Patiejunas, 2015; Mellor, 2013). This accumulation, combined with the constant aggregation of content from external sites such as newspapers, represents what was referred to as *information sediments* in the introduction to this chapter. The continuous increase in, and instant access to, these sediments can be used to make influence operations more believable by referencing existing material to make it seem like a mainstream opinion.

Past influence operations may have used radio channels under their own control, planting news items into local newspapers or TV stations or starting friendship associations (Abrams, 2016). What unites the above differentiators of social media from traditional media is the combination of not having any pre-post, human gatekeepers and the fact that the information presented forms part of a generally believed stream of information that the social media user deliberately seeks out.

2.4 Amplification and extended reach

A key element of the social media influence operations witnessed in relation to US elections and Ukraine is that the original content posted, and/or the points of view they contained, spread beyond the places of the original posts. This is a vital part of the type of influence operations discussed in this report. Spreading the message both makes the message more believable (Paul & Matthews, 2016) and has a larger chance of reaching people who will believe it.

The aforementioned targeting of relevant audiences through the content selection algorithms used by social media services is important here. However, to reach people who can help amplify the message the self-selection that users engage in by forming communities on social media[3] is equally important. Such communities may be online instances of existing organisations, or they may be ephemeral and organised around a topic. For instance, the sustained attack on specific women in the gaming world that is one of the three cases used in this report united a variety of people online under the hashtag #gamergate. This was a "*proto-social movement*" (Molden, 2015) which "*existed for years before it had a name: the same core players, the same harassment, the same abuse. The hashtag just put a name on this 'loosely organised mob' that attacked women in gaming*" (Anita Sarkeesian, interviewed in Valenti, 2015) .

An important element in these communities is emotions, which can be shorter lived, yet more intense, than offline: "*[the online] emotional regime is primarily a regime of emotional intensities, in which the amount of emotion matters*" (Serrano-Puche, 2016, p. 2). These emotions may in fact be what lead to people forming communities in the first place, emotions that can be sustained by the feedback affordances of social media: "*The more someone links to you, likes you, thumbs up your postings, and comments on them, etc., the higher you will be ranked and listed in the different SNS, news feeds, and tables of suggested links and readings*" (Svensson, 2013, p. 22).

The meaning of trolling

The act of posting inflammatory messages in social media online is often referred to as *trolling*. Currently this term is understood to relate to the aggression displayed in such messages. However, the original meaning of the term refers to a way of fishing (Donath, 1999; Herring, Job-Sluder, Scheckler, & Barab, 2002). Trolling (*dorging* in Norwegian) uses one or more fishing lines from a moving boat to catch fish.

Using the original understanding of this term makes more sense in connection with influence operations in social media. The perpetrators' goal is to fish for reactions and try to affect groups of people who go for the bait, i.e. any inflammatory social media posts.

The general disinhibition that many users exhibit online aids such efforts. Some people may take particular delight in attacking someone or being generally aggressive online as a result of negative personality traits (Buckels, Trapnell, & Paulhus, 2014; Craker & March, 2016), but even regular users can exhibit more hostile behaviour online (Cheng, Bernstein, Danescu-Niculescu-Mizil, & Leskovec, 2017), often as a result of what is known as the *online disinhibition effect* (Lapidot-Lefler & Barak, 2012; Suler, 2004, 2005; Udris, 2014), also discussed as the anonymity effect (Bjørnstad, 2019). At the same time, being part of a community that attacks others through online means strengthens the community as a whole and at the same time makes individuals feel good about themselves, encouraging further online attacks to sustain the feeling of community.

By exploiting existing divisions in a society, and target messages to connect with online communities' meaning making, you expand the base with willing helpers that will be unaware

[3] One example that has received attention recently is the *Vaccine Resistance Movement*, one of many groups that claim that certain vaccines will damage your child. See http://facebook.com/groups/VaccineResistanceMovement/

of your underlying intentions (Gioe, 2018; Morgan, 2017). Supporters of fringe views on a topic may also be more willing to spend the time it takes to disseminate information from an influence operation by actively spreading the message in different ways, using the affordances of social media mentioned in the previous section. This report suggests that the cold-war political term "useful idiots" (Oxford Dictionaries, 2018) can be used as an analytical lens to examine how influence operations can attract additional support. Until the 1990s this term was mostly applied to those who were judged to (possibly) be manipulated for political purposes by another (often communist) state. In this report the term is used in reference to someone's online persona as it manifests itself through their social media activities, use of the term does not imply any assumptions or judgements with regard to someone's offline actions or attitudes. The emphasis of both "useful" and "idiot" is therefore on discrete actions and not the whole, offline person. There may be more useful idiots than there are willing helpers for an influence operation, this helps an attacker quantitatively. As such social media users are not directly linked to the attacker they add respectability and believability to the attack, thereby helping the attacker qualitatively. In short, useful idiots can be force multipliers in a social media based influence operation (Bergh, A., 2018) and an attacker can harvest angry online exchanges as a form of free, crowdsourced microwork, similar to Amazon's "Mechanical Turk" marketplace (Fort, Adda, & Cohen, 2011; Irani, 2012), but without any financial costs. The energy expended in brief, angry online exchanges can thus be funnelled into a broader influence operation against a target.

The term "*affective news streams*" from the communication scholar Zizi Papacharissi neatly sums up what is happening in this intersection between community, fringe views and online aggressive behaviour that useful idiots inhabit. Affective news streams are "*news collaboratively constructed out of subjective experience, opinion, and emotion all sustained by and sustaining ambient news environments. We may understand affective news as the product of hybrid news values and ambient, always-on news environments. Affective news streams blend fact, opinion, and sentiment to the point where discerning one from the other is difficult, and doing so misses the point*" (Papacharissi, 2015, p. 7).

The generation, sharing (by people and algorithms) and consumption of news, i.e. information, becomes an important part of the meaning making of the community. Usher has suggested that "*[…] people are making conscious decisions to aid the circulation of certain content because they see it as a meaningful contribution to their ongoing conversations […] As they circulate this content […] they also help to frame the content"* (Usher, 2010). On a practical level the generation of new information online is often done through *remixing* existing content (Lessig, 2008). Such remixing is very common on the Internet and social media in general. In influence operations remixing can not only reduce costs in terms of time, money and skills required to generate content, it also allows one to distort or hijack the other side's narrative(s), as well as recycle (through sharing) old stories to start a new cycle of virality and/or discussion. This recycling is also done by algorithms that select information deemed to be relevant to what the user is currently reading. If a kidnapping case is currently in the news, old kidnapping stories will typically be recommended for the users reading about the current case. Thus old narratives can be re-energised through *trending topics* and similar automated story selections.

2.4.1 Amplifying by gaming the technology

Social media services use various algorithms to monitor, add, select and sort content when users add or search for it. Such algorithms are a potential resource to manipulate to promote one's own content and hide the other side's posts. For example, in the US anti-abortion clinics are using a range of Google provided tools to come on top in Google Maps searches: *"[anti-abortion clinics] that aren't already in the system are using [the] 'add a missing place' [functionality] to insert their own listings to the pool. Google itself offers a number of practices to improve a listing's rankings in the results, such as entering extensive business info, full contact info, photos, and responding to reviews"* (Marty, 2018). As such user submitted data is checked automatically without human intervention it is relatively easy to stop people from finding what they are actually searching for.

The type of attack described above uses valid content, but manipulates the automated routines that check and rank the content. Another variety of such attacks can be used to hide content that might be banned or could reveal the target of an influence operation. This can be done by using *adversarial machine learning*. Machine learning is the most common type of artificial intelligence used to recognise patterns in content, frequently used to analyse images to find out what is in the image. This feature is used to search for images that contain a particular object, or automatically add your friends name to a picture you upload to social media, etc. Adversarial machine learning is the attempt to negatively manipulate the machine learning in order to get a system to do something different than intended. Recent work has shown how such machine learning can be fooled by adding certain types of noise in the picture, where individual pixels are changed slightly, unnoticeable to the human eye, but the software believes it is something completely different. See figure 2.4 below. Other, recent research has shown the potential for tricking self-driving cars with "poisoned" street signs where advertising is interpreted as a stop sign or similar (Sitawarin, Bhagoji, Mosenia, Chiang, & Mittal, 2018).

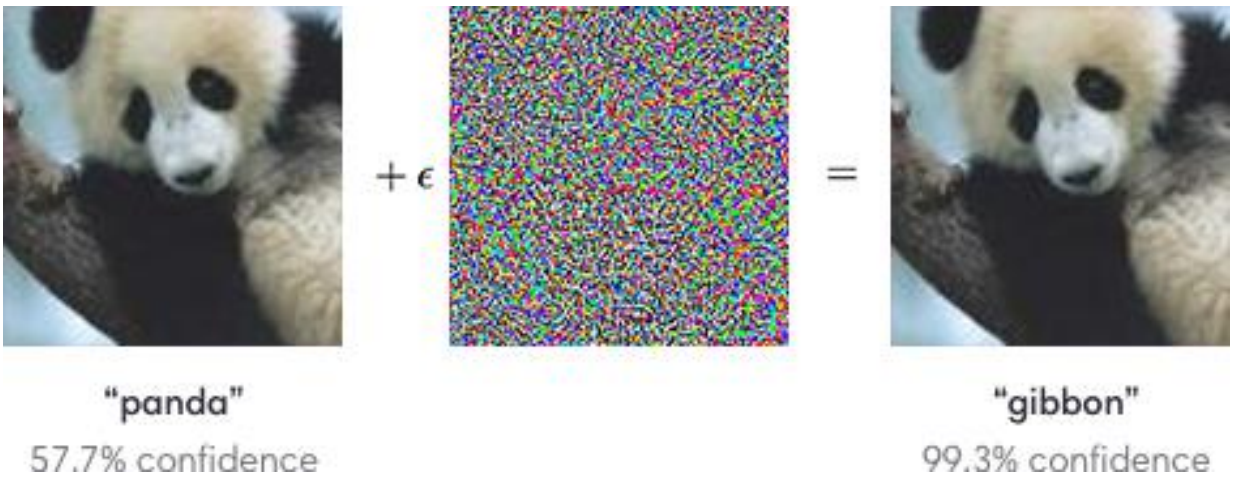


Figure 2.4 A picture is distorted and the machine learning software believes it is a gibbon. (Goodfellow et al. 2017)

The advantage of manipulating content selection, search and ordering is that most social media users assume that what they see online is a neutral view of what is going on. The impact of influence operations on people may be stronger when it seems that "everybody" in the groups they feel they belong to (such as conspiracy theorists) agrees with the message they are trying to push. This is already being taken advantage of by certain actors: *"Russia has embraced algorithm warfare information operations to disturb other nations' domestic stability. Russia cleverly uses others' algorithms against them, perhaps creating a whole new dimension to such warfare and suggesting a way smaller nations might manoeuvre in the new 'intelligentized' warfare era."* (Layton, 2018). This attempt at creating a false social consensus also promotes persuasion. This link is discussed in detail by Bjørnstad (2019).

2.4.2 Typology: Why and how of online trolls

The term troll is often used to describe Internet users who either attack someone personally online, or are posting offensive comment; mostly not as part of an influence operation. In reality there are wide varieties of aggressive behaviours online; these emerge from why someone is aggressive and how they choose to express that aggression. Trolling may be motivated by outside (extrinsic) or inside (intrinsic) forces; an example of the former could be because you work on an influence operation, the latter might be because you enjoy making others miserable.

The grid in figure 2.5 is based on the literature and events referenced in this report; it is not necessarily complete at this point in time. The purpose is to highlight what to the author appears to be the most common why-how linkages to see how different trolling activities are used in influence operations. These linkages are clustered in the highlighted part of the grid.

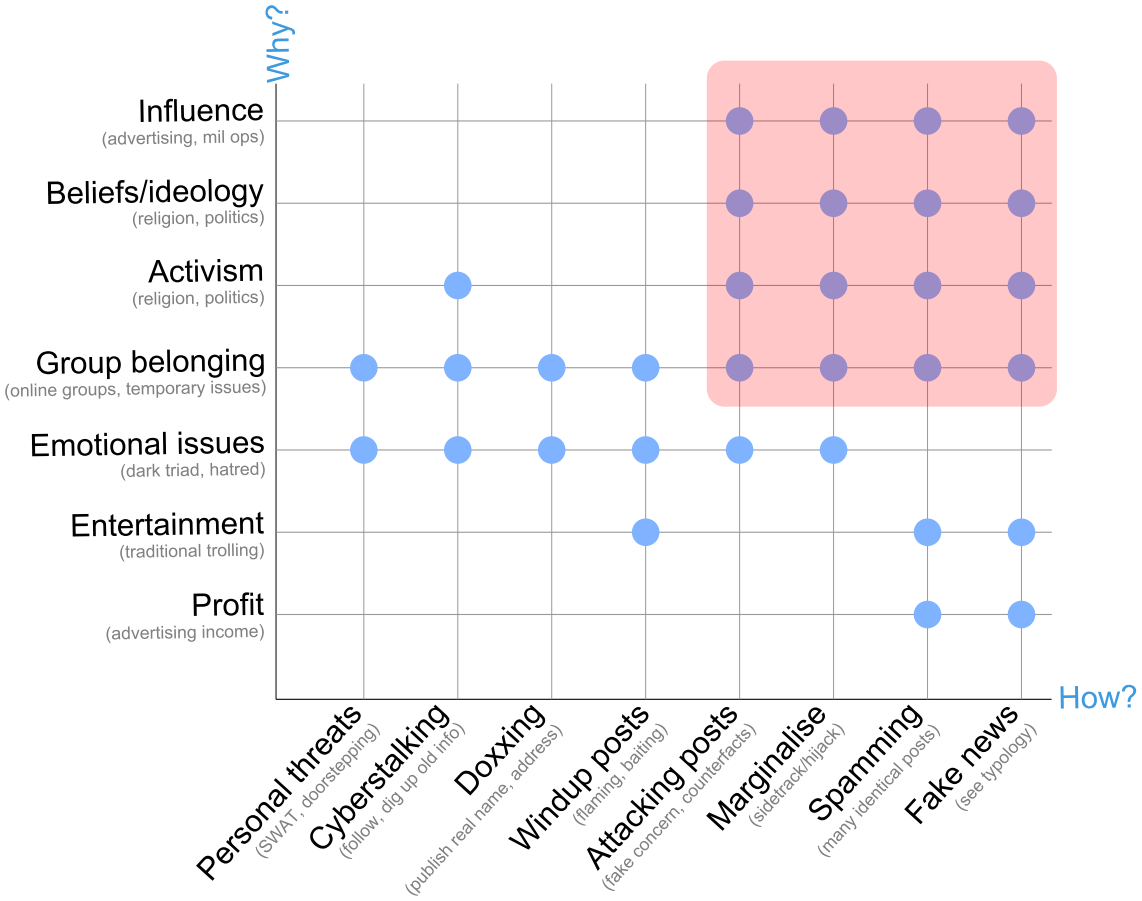


Figure 2.5 How and why online trolls act.

In the above figure the why axis motivations *influence* (the core of influence operations) and *beliefs/ideology*, *activism* or *group belonging* (motivations of the useful idiots discussed earlier), often resort to tactics focused on spreading the message wide (spamming or fake news) and attacking or marginalising opposing voices.

2.5 Online information sediments

An important part of our conceptual chain is how we perceive and discuss the vast amount of information that is continuously posted online. Numerous metaphors are used to try to describe the Internet and Internet based content, caused by our need to make the abstract nature of the Internet more understandable (Jamet, 2010; Tomaszewski, 2002). When we choose how to discuss the Internet it has "actual and meaningful consequences on the shape and perception of these technologies" (Markham, 2003, p. 1).

Thus, any metaphor of social media and social media content needs to reflect certain aspects related to influence operations that have taken place in social media:

1) It must emphasise the long term, cumulative approach of social media. The incessant accumulation and aggregation of information was discussed above as a differentiator between social and traditional media.

2) The metaphor should highlight the fact that despite information never disappearing, it will fade in and out of view. Past posts are generally pushed out of sight by new post. However, new posts about, or users searching for, a particular topic can cause algorithms to retrieve and display old information that is somehow linked to the new information.

3) Furthermore, how narratives emerge in social media can both affect, and be affected by, existing information. A large number of new posts that twists existing information in a particular way may change an established narrative. However, previous social media posts provide a frame that can have an effect on how new posts are perceived.

4) Finally, new influence operations, whether by design or not, don't start from a blank slate, they connect with existing posts that may already have influenced those reading the new posts. For example, research on online hate-speech suggests that *"despite the relatively short 'half-life' of antagonistic content towards Jews, once this temporary increase in online hate speech receded it left behind a new, higher baseline of online hate"* (M. Williams & Burnap, 2018, p. 6).

This report suggests that the metaphor of "online information sediments" provides us with tools to conceptualise the four issues above. Social media is in effect a river that carries along content that is active (i.e. in the stream) for a short time until it ends up as layers in a vast reservoir of information (i.e. *sediments*) as they age, similar to the debris of a real river. These sediments can force the river with new information to flow a particular way, but a strong river (large numbers of social media posts) can move sediments around, stir things up and create a new flow, as discussed above regarding remixing and reframing existing content.

In any influence operation using social media there will be an implicit long-term effect that contributes to building up the presence of the attacker's message or narrative in the online information sediment. This can also be seen as a further affordance of social media, other affordances were discussed in *2.3 Affordances of social media*. The presence of all these posts in the online information sediments could prime some people to believe the attacker's narrative(s). This would also tie in with the *sleeper effect*, this is when *"low credibility sources manifest greater persuasive impact with the passage of time"* (Paul & Matthews, 2016). It has even proved possible to implant entirely false memories by creating so-called autobiographical references that mix reality and fiction and present it as part of someone's past (Braun, Ellis, & Loftus, 2002), a long term campaign could also aim at this. Indeed, it has been suggested that long term approaches are key to Russian influence operations: *"The threat of Russian information campaigns is thus that they prepare the ground for future Russian action which would be directly counter to the interests of Europe and the West. By either undermining the will or support for deterrent measures, or sowing an entirely false impression that Russia is justified in its actions, Russia adjusts key variables in the security calculus determining the risk inherent in future assertive action against its neighbours"* (Giles, 2015).

In figure 2.6 is an example of how information sediments affected a search algorithm when Google's Polish search site was accessed through a computer located in Poland. When entering "Norway child" (in reference to the Norwegian Child Protection unit) the terms "stealing" and "kidnapping" is suggested by Google. BBC is also a suggestion, presumably a result of BBC News stories (on TV and website) about children being removed from parents in Norway (Whewell, 2018a, 2018b). When searching in Norwegian on a computer located in Norway the results were very different without any negative connotations.

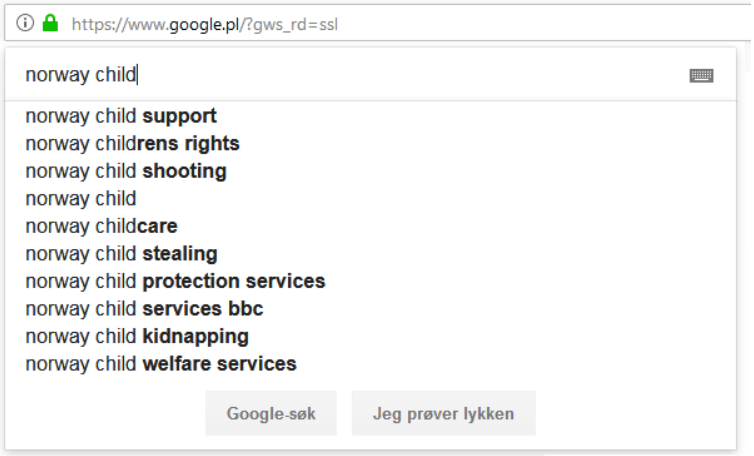


Figure 2.6 Google search suggestions when entering the first two words of "Norway child protection services".

When encountering an influence operation in social media it is thus beneficial to understand how to manage these online information sediments. This management is done through new social media posts or by using the algorithms that emphasise some types of posts while ignoring others. In other words, the information space should be moulded to benefit the defending state, rather than trying to build a dam to stop the flow of information.

2.5.1 Fake news

The emergence of a new term may obscure the fact that we are not dealing with something new, merely something that seems new. This is very much the case with so-called fake news. Made up news that is spread in the hope of influencing people goes back thousands of years. However, in the current context, what is new are those aspects that are unique to social media as discussed in *2.3.1 Social media differentiators* over. To recap these are direct access to social groups or societies around the world, low cost, no editors that check all content before it is published, the ability to blend into the everyday online content due to social media distribution and design and an extended reach when posts from influence operations are shared by others.

An important aspect on the current flood of *fake news* is that often the content represents information that many readers want to be true and that they actively seek out (Goolsby, 2019). This is because the stories reflect the world as they perceive it, and fake news becomes an important part of their (shared) meaning making. And one of the most important ways in which

such stories spread is through the recommendation algorithms of large social media services (Fiegerman, 2018; Levin, 2017; Lewis, 2018a; McKay, 2017), something that is knowingly manipulated by authors of these stories.

It is therefore generally more useful to discover the underlying reasons for fake news being spread than focus on its lack of facts or truthfulness when countering influence operations in social media. For many stories that present facts that can be checked (as opposed to opinions), it is not particularly difficult to find out if the facts are correct if you are inclined to discover the truth. Apart from reading reputable news outlets, there are long standing efforts by websites such as Snopes.com that debunk urban myths and fake stories online and relative newcomers such as Faktisk.no examine the correctness of claims in Norwegian media, whether traditional or online. This report would suggest that efforts such as these (ref. eg. Haciyakupoglu, Hui, Suguna, Leong, & Rahman, 2018; Pavleska, Školkay, Zankova, Ribeiro, & Bechmann, 2018; Sample, 2018) may be useful as a baseline for people to use in an argument, they are less useful in terms of countering the attacks because it allows the attacking side to set the narrative (Jankowicz, 2017). Furthermore, a working paper analysing the tagging of suspected fake news stories in social media found that this led to an implied truth effect. This meant that fake news stories that were not tagged were more likely to be assumed to be accurate (Pennycook & Rand, 2017). This suggests that when efforts are made to examine social media content, and such efforts are communicated to users, anything less than 100% success rate can backfire by making other stories more believable.

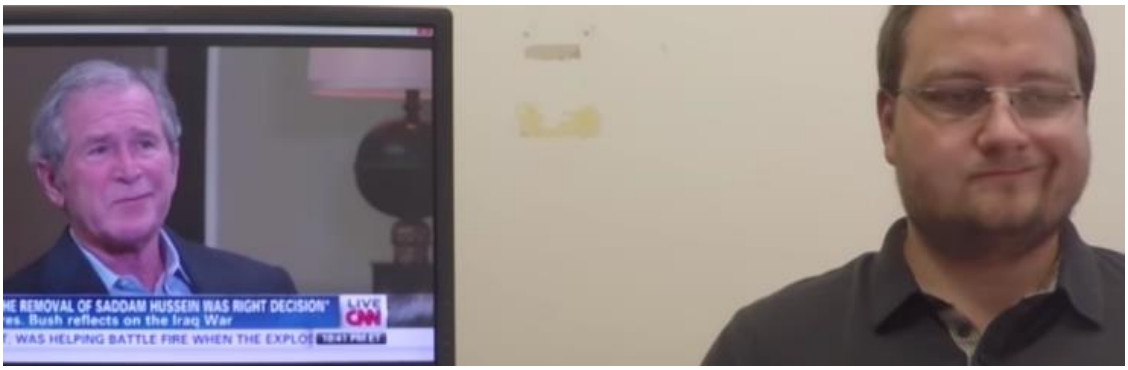


Figure 2.7 Deepfake video, George Bush's mouth follows the live actor (Matthias Niessner, 2016).

Current developments in the field of sharing and remixing visual information is perhaps the most worrying from a fake news point of view. So-called *deepfake* videos use an existing video but superimpose (thus replacing) someone else's face on top of the original face in such a way that the superimposed face perfectly mimics the original expressions throughout the video. The new development here is that these are now much more realistic and easier to create; one can download software and with some patience and not too much computer skills generate one's own deepfake videos (Quach, 2018) as figure 2.7 over shows.

2.5.2 Typology of fake information

The purpose of this typology is to explore the different shades of what has in recent years become known colloquially as fake news, as well as adjacent types of information. In the FFI report *Falske nyheter som sjanger* (Alme, 2019), fake news is defined very narrowly. The goal there is to look at why "pure" fake news uses the news format to connect with the target audience. In this report the aim is to examine the broad range of non-true information types that might be deployed in a campaign, hence the more inclusive approach to the term.

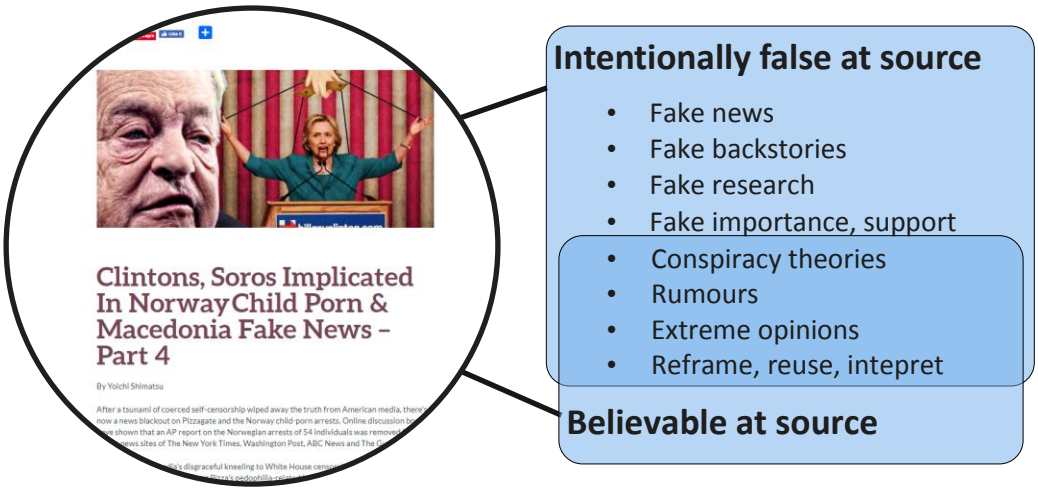


Figure 2.8 Types of information frequently referred to as fake news.[4]

This report suggests that the different types of information under the fake news umbrella term can be seen to fall into two broad, and partly overlapping, categories as shown in figure 2.8 above. The first is information that is deliberately created to mislead and misinform, and the creators know this. The second set is types of information that is believed to be true by the creator. An example of this could be a belief that the moon landings are faked, i.e. a conspiracy theory. This latter form of fake information can of course also be created deliberately by actors behind an influence operation. However, the fact that existing groups create these types of information outside an influence operation makes them a part of the "useful idiots" discussed in section *2.4*. These types of information are then created as part of these groups' everyday meaning making, making them more valuable to influence operations as they emerge from actors that have their own followers. Table 2.2 below has a short summary of the different fake information types as listed in figure 2.8 above.

[4] Screenshot from http://themillenniumreport.com/2016/12/.

Fake news	Planned and planted information related to real and current events, containing false information at the core, without everything in the story necessarily being incorrect. Such news will try to pass itself off as real news, either through design, placement or other news-esque features.
Fake backstories	Not news, but false claims about something in the past, related to current events. This may be claims about someone having gotten into a position of power through illegal or immoral means, or inversely, untrue claims about something done in the past by someone trying to obtain power through these claims.
Fake research	Research that is made up, misquoted, taken out of context, not real or based on old, outdated material. May be actual research, but badly carried out, biased or even real academics who falsify their research. Used to give an assertion added weight. An example would be the anti-vaccination movement who uses debunked research on MMR vaccination and autism (Wakefield et al., 1998).
Fake importance, fake support	This relates to the propagation of information, the falsehood here is how many regular people are interested in a topic and how important the topic is. The terms astroturfing is often used. It denotes seemingly grass root campaigns that are in fact staged by a central organiser who pay or encourage people to be involved and make this involvement seem bottom-up and uncontrolled. The Cambridge Analytica company would use this tactic by *"create[ing] content on the internet for them to find," he says. Those posts and blogs would have seemed organic and authentic, but they weren't"* (Christopher Wylie quoted in Kobie, 2018).
Conspiracy theories	Beliefs that an orchestrated campaign is hiding information from the public, such as vaccines being dangerous. Often relies on *Fake research* (see above).
Extreme opinions	What represents extreme opinions or not will vary according to context. In one country being pro-abortion may be seen as extreme, in another being against it might be equally extreme. The point here is that these opinions may generate questionable information to support their views. This information may then be believed by those who share the base opinion.
Rumours	Incorrect information that is filling an information vacuum, often in a fraught situation where people seek information.

Reframe, reuse, interpret	Information where the core details are true, but they are taken out of context, or promoted as more important than it is. For example, a study of more than 11,000 Twitter links from the Russian influence operation in the 2016 US election found that the majority referenced well-known US news outlets, and not fake news sites (Albright, 2018). This is also about interpretation; factual information may be used to support an assertion that goes beyond what can realistically be inferred from the information.

Table 2.2 Brief explanation of eight types of fake information.

In closing it should be noted that information that is, in the main, incorrect is a large part of everyday life for most people. It can be entertainment such as US and UK tabloid newspaper stories, a vivid example being the entirely false story about a British comedian entitled *Freddie Starr ate my hamster* (Daily Telegraph, 2012).

2.6 Attention, a contested space

Social media affordances can help social media based influence operations to initiate, amplify and/or extend the reach of their message(s); at the same time the content produced by the influence operation is added to the online information sediments discussed above. As with any social media content the goal at this stage is a) to be visible in the social media streams of targeted people and b) for that content to receive some attention from these people. This report suggests that without any attention from social media users, the social media influence operation must be considered a failure; this step in our conceptual chain leads into the meaning making processes. If social media users don't pay attention, i.e. read, listen to or view the content, it cannot possibly affect their meaning making.

Understanding the role of attention in social media is important for both sides and the attention economy theory is useful here. It is in part based on an early insight by the economist and political scientist Herbert Simon more than quarter of a century before social media emerged. Discussing organisations and the increase in available information he stated that "*What information consumes is rather obvious: it consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention [...]*" (Simon, 1971, pp. 40–41). Although not a mainstream economic theory (Terranova, 2012), in terms of social media based influence operations' need to reach online users to influence them, it has considerable explanatory potential.

The notion that *"there is something else that moves through the Net, flowing in the opposite direction from information, namely attention"* (M. H. Goldhaber, 1997) is important here. The need to compete for attention, as a scarce resource (M. Goldhaber, 2006), is not only driven by the vast amount of information generated online. It is also a result of the underlying funding model of most social media, namely advertising. What social media services is selling is their users' attention, or as the adage goes: *"If you are not paying for it, you're not the customer;*

you're the product being sold" (Andrew Lewis AKA 'blue_beetle', quoted in garson, 2017). This links back to the affordances of social media that are used by such influence operations to spread their message, these are developed to get more users and more of users' attention through engagement tactics such as *Like* buttons, tools for easily sharing content, highlighting news and posts that matches the user's profile, etc. However, sometimes such tools are used without any scrutiny from the user. For example, one study estimated that some 59% of links are shared without the sharing user themselves reading the page at the end of the link (Gabielkov et al., 2016, p. 5). This type of behaviour would help influence operations to spread their content to many more social media users, increasing the chances of encountering someone who ultimately pay attention to the content of the posts from the influence operation.[5]

In response to this competition for attention, a number of strategies are deployed by different social media and Internet actors outside of influence operations. However, these strategies are also, to varying degrees, seen in social media based influence operation as well. It is therefore worth summarising these briefly before discussing viral content.

Goldhaber, who developed the idea of the attention economy, suggested that *"this new economy is based on endless originality, or at least attempts at originality"* (M. H. Goldhaber, 1997). What can be seen is not exactly originality, but a continuous increase in exaggerations and hyperbole in the language and imagery used to elicit emotional reactions to grab someone's attention, even for a split second, usually by encouraging the user to click on a link to go to another webpage. These so-called *clickbait* tactics are used by a variety of actors, from traditional newspapers attempting to get more readers for a regular story (Tangen, 2017) or specialized outfits who make money from providing false news stories linked to current events, often the more controversial ones (Tynan, 2016; Kirby, 2016; Bergsaker & Bakken, 2018). Another tactic to gain attention is used by so-called content farms where (low) paid staff writes stories in response to market opportunities, for example looking at what Facebook users discuss (Dewey, 2015) or examining *"popular search terms from search engines; [and] which [advertising] keywords are currently being sought and how much is being paid for them"* (Napoli, 2014).

A third approach is to gain the attention of gatekeepers and content curators, i.e. well-known social media users, whether celebrities, so-called influencers or others who have a large following. These social media users may have several million followers (although usually a lot less) who are alerted about new posts. As mass media makes way for social media,[6] these users in some ways act as editors (albeit without any of the traditional responsibilities of editors) for their followers, selecting what information receives attention. As a researcher on social movements online explained *"[the] fracturing of publics, somewhat ironically, increases the importance of 'focusers' of attention, which can be institutions (media outlets), individual*

[5] As an example of the exponential distribution potential inherent in social media one can look at the scandal surrounding Facebook and the consultancy firm Cambridge Analytica. Although only "hundreds of thousands of users" used a Facebook quiz app, the app collected not only their own details but also that of all their contacts. In total this yielded 50 million profiles (Graham-Harrison & Cadwalladr, 2018).

[6] As an example, 32% of Norway's population read a newspaper daily (www.ssb.no/en/kultur-og-fritid/statistikker/medie/aar) whereas 80% use social media daily (www.ssb.no/en/statbank/table/11437/).

mediators of attention (and on social media, this includes prominent journalist-curators, such as Andy Carvin), celebrities […]" (Tufekci, 2013, p. 851). Their role is often linked to celebrity and not editorial responsibilities, thus Roseanne, a well-known American comedian with close to a million followers (Levine, 2018), has used her social media platform to promote conspiracy theories (Mandell, 2018). This has, in turn, earned her an interview on RT.com (Russia Today) (RT.com, 2015), the state sponsored media outlet that uses conspiracy theories, among other techniques, to sow doubts about democratic values and western leaders (Richter, 2017). Here one can see how attention to one topic can make you open to give some of that attention to a totally different topic, so an interest in comedy can expose you to conspiracy theories.

Figure 2.9 Cambridge Analytica presentation about their targeting capabilities (Concordia, 2016).

Finally, and most contentious, is the notion of psychological targeting, often through adverts, in social media and on the Internet in general. This is controversial for several reasons. Firstly there have been claims that it has been used to micro-target people in political campaigns by building up detailed psychological user profiles and then creating highly specialised content appealing to that group (Cadwalladr, 2018; Cadwalladr & Graham-Harrison, 2018; Graham-Harrison & Cadwalladr, 2018): The case of the UK based firm Cambridge Analytica and the US election was partly based on data about individuals that have been obtained through Facebook by subterfuge (Dwoskin & Romm, 2018; Lewis, 2018b; Parakilas, 2018; Temperton, 2018), which is highly problematic from a privacy angle. Finally, there is considerable doubts as to the efficacy of the methods, how much is exaggerations by commercial companies to get customers, and how much is actually working (Armstrong, 2018; Doward & Gibbs, 2017). Sidestepping the discussion about the usefulness of specific approaches, similar tactics will undoubtedly be attempted in future influence operations.

The different tactics outlined here are often used to sell products through advertising (Braun et al., 2002; Einstein, 2016; A. M. Kaplan & Haenlein, 2011), however they may also be useful to

social media based influence operations. This can be in the first phase where the campaign's own posts need attention to enter into the meaning making processes. It can also be in the secondary phase where the useful idiots ecosystem is facilitating further spreading of the content as well as creating their own, supporting content; in both cases increasing the chances of the content reaching people that will give it some attention.

2.6.1 Understanding viral content

The holy grail of Internet content producers who want large scale attention, whether for influence operations, to sell a product, or getting votes, is that the content goes *viral*. This term, first used in 1996, has been described as:

> word of mount (people recommending something to friends, family or colleagues)
> + exponential growth (social media affordances facilitating faster and wider spread)
> = viral marketing (A. M. Kaplan & Haenlein, 2011).

The key here is the exponential growth; the act of forwarding and sharing news in social media is simply how many people get relevant news (Stelter, 2008) and is not in itself a sign of virality. Viral exposure, and any effects from it, are free to the content creator through a combination of social media infrastructure (storage, sharing mechanisms, etc.) and the free labour that social media users engage in (Terranova, 2000). Given this, and the potentially massive reach this results in, virality can be highly desirable for broad influence operations. Added to the quantitative dimension is also a qualitative dimension. As virality is a form of electronic word of mouth, those exposed to the viral message may trust the content more, since it has been shared by someone you know.

This report suggests that virality depends on 1) the content is appealing to certain type of social media users, and 2) that users are willing to take time to perform certain actions. These two parts are themselves made up of two parts. 1a) is the content itself and 1b) is the perceived importance or popularity of the content, i.e. how many have liked it, retweeted it, etc. The users' actions are 2a) the willingness to engage with the content and 2b), spend time sharing it.

The content from influence operations that has been found in social media recently is usually aimed at sowing distrust, often appealing to existing fault lines in the society targeted (Brooks, 2017; Morgan, 2017; Shapiro, Leslie, 2017) or when used alongside a kinetic conflict, to put across a narrative that suits the attacker, such as Crimea being Russian as discussed earlier in *1.3.1 Example 1: Annexation of Crimea*. Often it relies on simple visuals or angry content aimed at triggering emotions to achieve a response.

The targeting that is done to achieve virality is more rudimentary than the type of psychological targeting discussed above. This lies in the nature of virality, you want large audiences, not small pinpointed ones. One may still use target audience analysis of some sort based on geographical location, membership of online groups, subscription to certain Twitter feeds, etc. (Keating, Schaul, & Shapiro, 2017).

When it comes to making the content seem important or popular, those who have tried to influence people, whether in social media based influence operation or in other contexts, generally resort to automated means. This includes generating fake engagement data, for instance retweeting tweets to boost them in the automated rankings, this is frequently done using bots (cf. e.g. Howard & Kollanyi, 2016; Stella, Ferrara, & De Domenico, 2018; Arnaudo, 2017; Mezzofiore, 2018; Bessi & Ferrara, 2016; Shao, Ciampaglia, Varol, Flammini, & Menczer, 2017). Sometimes bots are partly controlled by people, either to make them less easy to detect (Grimme, Preuss, Adam, & Trautmann, 2017) or to make followers more efficient in spreading information (Timberg, 2017). Bots have also been found to focus on particular topics, and then to be repurposed for a new topic, thus building on an existing base of followers (Howard & Kollanyi, 2016, p. 2). It is also possible to use commercial services to inflate social media statistics such as the number of clicks, retweets or followers (Deahl, 2017; Reinstein, 2018; Cresci, Di Pietro, Petrocchi, Spognardi, & Tesconi, 2015). This is typically done by using low paid people to manually click on interaction tools, such as a *Share* or *Like* button that increases the counters used to determine how popular a social media post is.

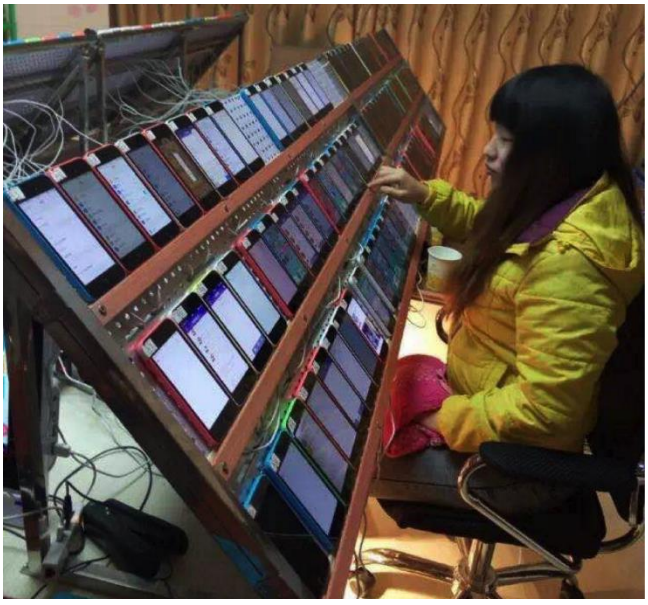


Figure 2.10 Worker in a click farm in China using dozens of smartphones (Cheaib, 2017).

The exact scale of the problem of fake popularity is difficult to estimate, however when Twitter, after considerable bad publicity about Russian use of Twitter to spread propaganda, at one point removed 70 million accounts in two months (Timberg & Dwoskin, 2018), their quarterly base of monthly active users dropped by 9 million (Henshel, 2018). More interesting perhaps is the effect this purge had on the number of followers high profile users had, in one day the president of Rwanda lost a third of his followers, whereas former US president Obama lost 3 million in the same purge (Jacobs, 2018). How many of these followers were commercially purchased fake accounts or fake likes generated by commercial operations and how many were done by software bots is of course impossible to say.

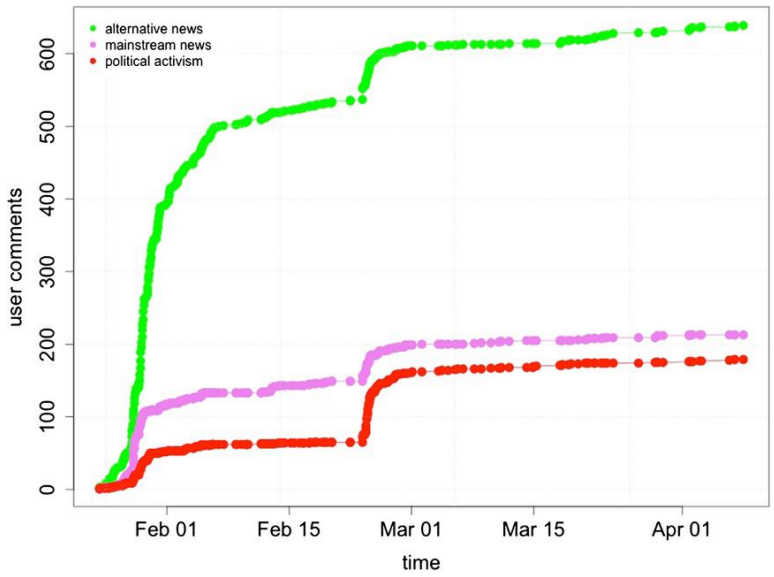


Figure 2.11 Cumulative number of comments per day on different types of Facebook posts (Bessi, Scala, Rossi, Zhang, & Quattrociocchi, 2014, p. 10).

Finally, virality in the end depends on content actually being passed on, and thus some work done by the user. There are many reasons why a social media message is shared but emotional reactions seem to play a prominent role here. Research suggests that *"emotionally charged Twitter messages tend to be retweeted more often and more quickly compared to neutral ones"* (Stieglitz & Dang-Xuan, 2013). These findings are supported by research that shows how users expressing *moral emotion* around topics like abortion rights, are more likely to share content that they deem morally important (Brady, Wills, Jost, Tucker, & Bavel, 2017). A review of research on sharing and emotions also found considerable support for this idea, suggesting that sharing information can be about both sharing emotions and information at the same time (Dafonte-Gómez, 2018, p. 2142). In terms of influence operations in social media, consumers of fake news, as discussed above, were found to be have more online interactions with the posts they read (in an Italian sample), as shown in figure 2.11 over (Bessi et al., 2014).

2.7 Individuals' and groups' meaning making processes

This is the key link in the conceptual chain that this report has developed (see chapter 2 for the overview) to help us better understand the type of influence operations in social media discussed herein. The underlying idea is that any (in)actions an influence operation wants to trigger in a target population need to be activated by the incorporation of information from the campaign into the targets' understanding of the world. This does not need to, in fact probably should not, attempt to challenge someone's world view head on. Neither does it need to be permanent, a short, even a few seconds response can be enough to fulfil the objectives of an influence operation on social media. The approaches that have been found in the influence operations discussed here are generally about making existing convictions stronger or sowing doubts about common, existing discourses (cf. e.g. Morgan, 2017; Richter, 2017).

2.7.1 Examples of influence operations content and links to meaning making

It can be useful at this point to examine some simple, randomly chosen examples of Russian influence activities (not necessarily full blown influence operations) before continuing the examination of what meaning making is, and some common elements that aids social media based influence operations.

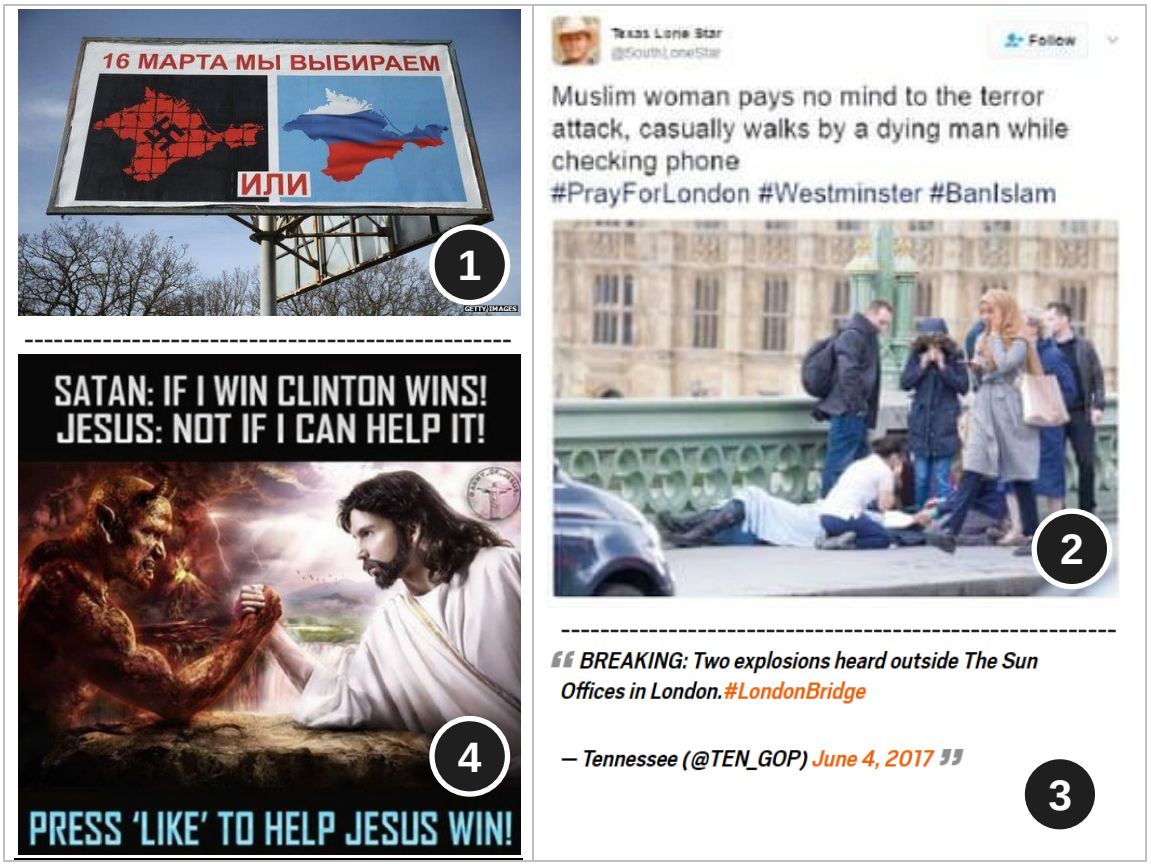


Figure 2.12 Four examples of Russian content 1: Poster from Crimea before the illegitimate referendum (BBC Monitoring, 2014), 2: Tweet posted after terror attack in London (Hern, 2017c), 3: Tweet posted after same attack (Mann, 2017) and 4: paid advert from Facebook (Permanent Select Committee on Intelligence, 2018).

Example 1 in figure 2.12 above is an attempt to use memories of the 2nd World War, when parts of Ukraine collaborated with the German occupation to establish a link between the Ukrainian government and (neo-)Nazism. Examples 2 and 3 are attempts at increasing existing divisions in the UK between the majority population and the minority Muslim population. In example 2 this is done by reframing a photo and claiming it shows a Muslim woman ignoring a terrorist victim. In fact her distress is clearly seen in other photos by the same photographer (Evon & Mikkelson, 2017), whereas example 3 simply posts false information aimed at increasing fear. The latter example was quoted in UK mainstream newspapers (Mann, 2017), one search found that this transfer from Russian controlled Twitter accounts to mainstream press happened at least 80

times (Hern et al., 2017). Finally, example 4 is part of an advert purchased through Facebooks normal sales channels and shown to targeted users in the US using standard Facebook demographic controls (Permanent Select Committee on Intelligence, 2018); representing a simpler version of the psychological targeting discussed in section 2.6.

2.7.2 Meaning making

When information such as these four examples gets the attention of a social media user, they become part of that person's meaning making processes (briefly introduced in section 1.4.1), the practices we all engage in when trying to make sense of the world. Such meaning making is often a collective creation, where categorisation helps people make sense of the world (Dobbin, 2009), to "*produce identification, commonality, connectedness and groupness*" (i.e. different forms of 'togetherness') in social media (Leppänen, Kytölä, Jousmäki, Peuronen, & Westinen, 2013, p. 1). On a broader level one may say that someone's cultural backgrounds are the result of meaning making (Spillman, 2002, p. 4). This report will discuss meaning making on the sociological level, looking at online group(ing)s. For a more in-depth examination of issues raised in this section please refer to *"Understanding influence in a defense context: A review of relevant research from the field of psychology"* (Bjørnstad, 2019).

In the type of influence operations examined in this report, the goal has typically been to make social media users believe some implicit or explicit statement that is made. For such information to enter into someone's meaning making processes the information must be convincing.[7] This report has discussed the use of automated bots to make content from influence operations more visible (see 2.6.1), this can also make it more believable. Research examining how people accept content suggest that retweets play an important role (Kim, 2018; H. Lee & Oh, 2017; Morris, Counts, Roseway, Hoff, & Schwarz, 2012; Oh, Agrawal, & Rao, 2013), as does information that claims some form of evidence and has an appearance of objectivity (Paul & Matthews, 2016). Users of social media sometimes evaluate credibility based on the amount of time spent interacting with audiences (Jahng & Littau, 2016), or user names (Morris et al., 2012). Repetition of information, even when fake or implausible, (Paul & Matthews, 2016; Pennycook, Cannon, & Rand, 2017 (not peer reviewed)) also makes users evaluate it as true more frequently.

These findings all point to the importance of the online information sediments in influence operations: The more supporting information a user can find, the more believable something is, thus the more aggregated and accumulated content there exists, the more someone's message or narrative is strengthened.

There are numerous mechanisms involved in meaning making; two that are particularly helpful to influence operations in social media are filtering and categorisation. The **filtering** of

[7] There are several caveats here. It is of course also possible to create and distribute information that you do not want to be believed. For instance in a so-called *false flag* approach one could imagine that nation A created non-believable information that looked like it came from nation B, making it seem like they were behind an information operation. Furthermore, even if false information is correctly identified as incorrect it can affect someone's meaning making, but then not necessarily in the direction sought by the information operation.

information, leading to so called echo chambers where everyone is in agreement (Krasodomski-Jones, 2017; H. T. P. Williams, McMurray, Kurz, & Hugo Lambert, 2015) is often linked to the algorithmic selection of information (Dias, 2014; Treré, 2016). However, a study on how users react to news that does not match their ideological background found that *"compared with algorithmic ranking, individuals' choices played a stronger role in limiting exposure to cross-cutting content."* (Bakshy et al., 2015, p. 1130). Filtering information is a form of confirmation bias, whereby new information is only incorporated into meaning making when they confirm existing beliefs. For instance, terrorist attacks done by Muslims were found to receive much more (traditional) media coverage in the USA than terror attacks by non-Muslims (Kearns, Betus, & Lemieux, 2017).

Categorisation of others, whether individuals or groups, is a key element of meaning making. There is considerable research on the concept of in and out-groups, the dividing up of people into "*those who are like me*" and those who are not (Barth, 1969; see also Bjørnstad, 2019 for more on the psychological processes of social media and in-groups). Such groups might be something one choose to belong to (joining a club) or innate (belonging to an ethnic group). Typically, the in-group will discriminate against, or make negative judgements of, the out-group and its members through stereotyping. In figure 2.13 there is an example of this. Two Twitter users, for and against Donald Trump, each accuse "the out-group" of propaganda or stifling free speech. Online these tendencies can be exaggerated as it is much easier to form an in-group (which also becomes an echo-chamber), given the worldwide and instant reach to others like you without the need to invest a lot of time and in great measures helped by algorithms and social media affordances. These purely online in-groups can be ad-hoc, for instance people who self-identify with certain view points in a Twitter conversation, or it can be long lasting and fairly formal through the membership of a closed Facebook group for instance. Often groups are set up specifically to be anti-something, for example the *Stand up for Sweden* anti-immigration group (Merrill & Åkerlund, 2018), thus further honing the us and them group feeling.

Figure 2.13 Two tweets responding to a Donald Trump tweet.

Finally, the elements discussed here, increased believability of social media content through repetition in different ways and the filtering of information and the formation of in-groups, imply that it is not required to do a very detailed analysis to find the relevant audience for your

content. If the purpose of the influence operation is to create negative views of some target group one can probably rely on self-selection and algorithmic selection. Besides, filtering and categorising information and people also aids existing narratives. If a narrative about Norwegian child protection services "*acting like Nazis*" (Christopoulou, 2018) is accepted by someone, then new information is likely to be processed to enhance, rather than counter, that narrative.

Narratives in social media based influence operations

The narrative as a unifying, overarching story can be seen as a core element of meaning making that *"[…] make sense of the world. They put things in their place according to our experience and then tell us what to do."* (Lucas & Nimmo, 2015). The concept is explored in social sciences (Czarniawska, 2004), strategic communications (Hagen & Søgaard, 2013, p. 10) and recent influence operations (e.g. Faizullaev & Cornut, 2017; Hutchings & Szostek, 2015; Biersack & O'lear, 2014). This report proposes that when analysing narratives around an influence operation in social media one could also see if any of the following sub-types characterises the process of creating the narrative.

The **algorithimic narrative** (Rourke, 2015) is generated by social media automated software routines that select news, generate information feeds and search for information by matching what the user already believes.

The **censored narrative** emerges as a result of by the attacking side censoring online information for its own population. In particular, their narrative might be strengthened by letting through certain external sources that support the narrative in their favour.

A **defensive narrative** is focused on responding to a narrative from the attacker, this can leave it prone to inconsistencies and difficult to maintain.

The **hijacked narrative** emerges when the attacker uses your own narrative and turns it against you by reframing it with new, possibly false, information.

A **deconstructed narrative** occurs when an influence operation explains away individual parts of a narrative about their actions. For example, the attacker can provide individual excuses for different elements in the narrative, so "invasion of X was humanitarian", "incident Y was a misunderstanding", etc. to undermine the overall narrative.

2.8 Encourage alternate individual or group (in)actions – is it possible?

This report has examined how influence operations in social media can utilise affordances of social media and certain behaviours of unwitting helpers to spread content to large audiences

around the world. This content, whether created by the influence operation or by sympathisers, enters into social media streams before becoming a part of the vast online information sediments. Along the way information from, and related to, an influence operation competes for attention just like any other social media content. If attention is received the information may then have an effect by entering into individuals' and groups' process of meaning making, i.e. making sense of the world around us.

It has been suggested that turning attention into actions may be a goal of attention seeking and that this *"[…] comes with the territory. That is part of the power that goes with having attention […]"* (M. H. Goldhaber, 1997). However, is this possible in social media based influence operations?

The first question would be whether people initiate actions based on information distributed through social media? The answer here is a clear yes, some examples include

- Parents who stopped vaccinating their children against (for instance) measles as a result of refuted and retracted research (Wakefield et al., 1998) that is fuelling online and social media conspiracy theories (Belluz, 2017; Dunn, Leask, Zhou, Mandl, & Coiera, 2015).

- At least 25 people have been killed in individual mob attacks in India, in response to fake news about child abductions spread in private messages on WhatsApp (Allana, 2017; Biswas, 2018). This problem has caused WhatsApp to restrict sharing (as opposed to writing new) messages (Cellan-Jones, 2019), a move that highlights how social media affordances play a role in users behaviour, it is not merely an online version of mouth-to-mouth rumour spreading.

- In Sri Lanka communal riots broke out after incitements to violence were spread via Facebook posts (Safi & Perera, 2018). Facebook's global moderators (discussed earlier) who did not know the local language took a long time to respond to reported posts and in one instance apparently responded that a post saying *"Kill all Muslims, don't even let an infant of the dogs escape"* did not violate community standards (Safi, 2018).

These examples show that influencing people to perform actions through a process that starts with giving attention to (untrue) information spread online is, in principle, possible. How specific the actions undertaken in response to such information can be is impossible to know for sure. There is also considerable debate about whether it is possible to persuade someone to change their mind through (social) media. These debates are usually related to elections and politics. Some research suggests that there is little or no change in target audiences' views. Reasons cited include few people actually being exposed to the information (Guess, Nyhan, & Reifler, 2018) and that there is little evidence people are influenced by, for instance, political campaigns of any type (Broockman & Green, 2014; Kalla & Broockman, 2018). It has also been suggested that on a neural level, the brain resists challenging information in the same area that links to identity (J. T. Kaplan, Gimbel, & Harris, 2016). In other words, it may be that information we don't already believe in requires such a deep change that it affects our identity,

which represents deeply held beliefs, making such changes in meaning making difficult to achieve. However, the examples discussed here are all from the USA which in the period covered by this research has experienced a strong polarisation in its domestic politics. This may mean that individuals researched would have had to challenge quite strong beliefs for any influence to be evident.

This report argues that the influence operations discussed here seem to have gained traction, i.e. their content being shared, remixed, reused and referenced in other media, when they enhanced existing views or divisions among their target audiences rather than trying to change their beliefs. Nudging existing groupings further in a direction they were already headed can be sufficient. Gamergate united existing misogynistic tendencies and the Russian interference in US elections focused on inflaming existing anti-*something* sentiments, e.g. Hilary Clinton, immigrants or police violence. The conceptual chain developed here suggests that social media based influence operations are useful to the attacker if it succeeds in getting targets to undertake alternate (in)actions. However, alternate does not imply "the opposite". What it means is actions that would not be done without the influence operation entering into targets' attention and meaning making processes. A simple example would be to go out and vote for someone you already support and not just express that support in social media posts. Or, and just as important, the non-action of **not** voting for someone you usually vote for.

3 Conclusion

This report has provided a socio-technical exploration of recent social media based influence operations, for example the Russian attempts at influencing the 2016 US presidential election and their use of social media in hybrid warfare in the Ukraine. The goal has been to expand our understanding of how such influence operations may have an effect and the processes that can lead to these potential effects. This has been done by examining these influence operations in conjunction with the wider literature on social processes, including online behaviour, meaning making and technical aspects of social media, for instance algorithms (that is, automated software procedures) used to select content to display and devices used to access that content.

The main contribution of this report has been to develop a conceptual chain of tools and arenas that are connected through different activities. The end point of this conceptual chain was also the departure point for the exploration of influence operations in social media. The point was that someone (but not necessarily everyone) targeted by a social media influence operation must a) in some way perform one or more actions that benefit the actor (e.g. a state) behind the influence operation and b) this action would not be done without the target encountering information stemming from the influence operation. If the targeted people changed their opinions to match the sentiments in the influence operations, but all outward actions were the same as before, then the actor running the influence operation would not benefit. If individuals

in state A were persuaded by information from state B claiming that sanctions against state B should be lifted, yet still voted for parties that promised to keep up the sanctions, then there is no useful effect for state B.

Working backwards from this premise, and exploring the unique aspects of social media and the influence operations that take place there, the following conceptual chain was developed to provide explanatory potential for how influence operations may have an effect (see chapter 2):

> A planned influence operation is executed by active operators that rely on the affordances (characteristics that facilitate certain activities) of social media. These affordances aid the amplification and reach of the influence operation so that the content created for the influence operation is spread widely and is added to the continuously aggregated and accumulated content stored by social media services. This vast content collection is referred to as the *online information sediments*, a metaphor that emphasises the long term, cumulative approach of social media where information never disappears but will fade in and out of view depending on what a user is interested in, what they search for, etc. New content added to social media does therefore not work on blank slate, it can both be affected by or affect the online information sediments. New content is *affected* because existing posts will provide material for framing and understanding the new posts. Conversely, new posts can *affect* existing content by providing a different way of interpreting and understanding existing information. Either way, information from influence operations is fighting for individuals' or groups' attention in social media to enter into and manipulate their meaning making processes so as to get the users to do, or not do, something that is beneficial to the actors behind the influence operation.

3.1 Future research

Given the potential benefits and comparatively low costs of using social media for influence operations, it is unlikely that the threat will go away in the near future. For example, before the 2018 mid-term elections in the US, attempts at influencing the election in a manner similar to the 2016 election was discovered. However, this time the accounts were disguised better, so one could not clearly establish the provenance of the account holders (Fandos & Roose, 2018). Another example of social media and fake news tactics not going away is EU's allegation that the seizure of three Ukrainian ships in the Azov Sea in 2018 was preceded by a year-long disinformation campaign by Russia (Boffey, 2018).

We can assume that tactics and tools will advance at the same rate of improvement as general Internet technologies. It has already been discussed how deepfake videos can be used to relatively easy map a different face onto an existing video and that it is increasingly difficult for viewers to detect. Over time one can envisage entirely synthetic, yet believable, fake online presences are achieved. In the arts this already happens. Hatsune Miku is a vocaloid, a synthetic person that is computer generated, but tours and performs; her fans write songs and votes are organised to select which songs "she" should record (Prior, 2018, p. 139).

Figure 3.1 Hatsune Miku in concert © C. Fountainstand (flickr.com/photos/9296771@N06/15426549459)

To build on the findings in this report and move towards more practical issues it will be useful to undertake further research in three interlinked areas, 1) detection and situational awareness; 2) content creation and delivery and 3) what aspects of the two aforementioned items can be automated through the use of software. Decisions such as whether to respond to influence operations detected, and if so, with what means, is beyond the scope of this report.

3.1.1 Detection and situational awareness

In an interview with staff in the Obama administration regarding Russian influence operations in the 2016 US elections it was pointed out that *"What we needed, and still don't have, is an analytic cell that sees the full scope of Russian activity. Our inability to put the full picture together in real time was a major part of why this was missed"* (O'Sullivan, Devine, & Griffin, 2018). Such ongoing and automated analysis is the first, and perhaps most important, step in handling influence operations in social media. One approach here could be a so-called *information environment assessment* that examines content from one's own organisation as well as third party and adversary content (Goolsby & Carley, 2019). A basic example of this was done in the Trident Juncture 2018 NATO exercise which collected and analysed tweets related to the exercise. Such work would require further research into how new or existing software could be used for such a task. Research could also explore divisive topics on the national level and issues that can affect a nation state's standing at the international level; these are likely to be exploited by an influence operation in social media.

A key issue with regard to situational awareness in social media is that different users experience different situations on the same social media service, at the same time. The vast amount of online information sediments will be endlessly and ceaselessly arranged and rearranged, selected or deselected, sorted and highlighted according to each user's accumulated

online history. This will be done instantly by different proprietary algorithms that are continuously updated without user input (Manthorpe, 2018; Wallaroo Media, 2014) and whose decision-making logic is buried in machine learning software which may or may not be biased (Birkbak & Carlsen, 2016; Caliskan, Bryson, & Narayanan, 2017).

When talking about situational awareness one should therefore not aim to see a single, summarised overview of all posts from a particular selection of social media. That would simply be a generic summary without any contextual information about who sees what. What is needed is to be able to see how other social media users will experience the content of the influence operation, and how this is framed by related content from other actors and data algorithmically selected from the online information sediments. It would therefore be beneficial to research the concept of personas from software development (Cooper, Reimann, Cronin, & Cooper, 2007; Blomkvist, 2002; Massanari, 2010; Pruitt & Adlin, 2006). Personas are about developing *"an archetypical representation of real or potential users. […] The persona represents patterns of users' behaviour, goals and motives, compiled in a fictional description of a single individual."* (Blomkvist, 2002, p. 1). Using personas would give a situational awareness related to categories of people, for instance the alt-right in the USA or Putin supporters in Russia without this infringing on anyone's privacy as the personas are not linked to any real people.

3.1.2 Content creation and delivery

At the core of social media influence operations is content that is distributed to users. To improve the ability to counter possible influence operations it would be useful to research if it is possible to create content on an ongoing basis by contributing to existing online services. This will provide searchable data on topics that can increase the nation's positive profile in the online information sediments. Related to this one should examine if one can *crowd out* content from an influence operation as well as the possibility of *entertaining and distracting* with own content.

"It's not nearly enough to create a good piece of content. You have to understand how content spreads across the web" (Jonathan Perelman at Buzzfeed quoted in Himler, 2013). This issue has been central to the discussion above, in particular sections 2.4 and 2.6.1. Linked to distribution is discoverability, i.e. how easy is it for a user who searches for something to come across the content that has been created? Research on the following areas would be useful here. Advertising (paid display) opportunities to bypass social media services own content selection. Moderation rules that different services use, this is so that content created to counter an influence operation is not censored. How do different content selection algorithms work? Facebook for example have implemented 58 changes since 2006 (Wallaroo Media, 2014). At one point one of these changes removed local, independent pro-democracy news from users' feeds in Cambodia (Kozlowska & Kozlowska, 2018; Paviour, 2017). Finally, what financial incentives are there for other actors to contribute in social media influence operations? This includes those who deal in content to earn money on advertising (Kirby, 2016; Wendling, 2018), or services, such as click farms, used to spread content or making it seem more popular (Wen, Cao, Shen, & Liu, 2018; Cresci, Pietro, Petrocchi, Spognardi, & Tesconi, 2014).

In terms of what audience to deliver the content to, two proposals that may be at odds with recent (strategic) communication perspectives could benefit from experimental research. 1) do not use target audience analysis (Tatham, 2015; Tunnicliffe & Tatham, 2017) and 2) do not focus on overarching (counter-)narratives (Lucas & Nimmo, 2015, p. 16). It can be surmised that social media algorithms that select relevant content for users will, in conjunction with users' self-selection through online group memberships, provide good enough (but not necessarily perfect) targeting. The same mechanisms will, this report suggests, help to link individual elements from the online information sediments, so as to make a convincing *algorithmic narrative* (Rourke, 2015) (see 2.4.1 Amplifying by gaming the technology). These suggestions could be tested out and compared in experimental settings to examine how one can manage algorithms and the vast quantity of online information sediments.

3.1.3 Automation through software

This report has explored the use of software automation in term of algorithms and so-called bots, in some cases discussed their shortcomings, at other times examined how they can be used to improve efficiency. However, there should always be human control over both content and content distribution. In a review of the history of Russian online influence operations Gilles points out that *"[in 2011 a] large array of pre-positioned Twitterbots, and sporadic but highly targeted DDoS attacks, were combined with old-fashioned dirty tricks against opposition leadership figures to attempt to defuse and discredit the protest movement. Examination of the results appears to have led to the conclusion that automated systems are simply not sufficient, and dominating mass consciousness online requires the engagement of actual humans."* (Giles, 2015). The so-called troll factory in St. Petersburg, The Internet Research Agency (@DFRLab, 2018; Lister, Sciutto, & Ilyushina, 2017) is a partial result of this. Similar approaches could be explored to see how computerised resources can be used to augment human output.

References

Abrams, S. (2016). Beyond Propaganda: Soviet Active Measures in Putin's Russia. Connections, 15(1), 5–31.

Adams, D. B., Brown, A., & Tario, C. (2009). Military Influence Operations: Review of the Consumer Psychology Literature. 39.

Albright, J. (2018, February 15). Trolls on Twitter: How Mainstream and Local News Outlets Were Used to Drive a Polarized News Agenda. Retrieved 13 February 2019, from Medium website: https://medium.com/berkman-klein-center/trolls-on-twitter-how-mainstream-and-local-news-outlets-were-used-to-drive-a-polarized-news-agenda-e8b514e4a37a

Allana, A. (2017, June 21). WhatsApp, Crowds and Power in India—The New York Times [News]. Retrieved 26 June 2017, from New York Times website: https://www.nytimes.com/2017/06/21/opinion/whatsapp-crowds-and-power-in-india.html

Allen, T. S., & Moore, A. J. (2018). Victory without Casualties: Russia's Information Operations. Parameters, 48(1), 59–71.

Alme, V. (2019). Falske nyheter som sjanger (FFI-Rapport No. 19/00660; p. 36). Kjeller, Norway: FFI.

Armstrong, S. (2018, March 22). Cambridge Analytica's 'mindfuck tool' could be totally useless [News]. Retrieved 22 March 2018, from WIRED UK website: http://www.wired.co.uk/article/cambridge-analytica-facebook-psychographics

Arnaudo, D. (2017). Computational Propaganda in Brazil: Social Bots During Elections. Computational Propaganda Project Working Paper Series.

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. Science, 348(6239), 1130–1132.

Bandaru, K., & Patiejunas, K. (2015, May 4). Under the hood: Facebook's cold storage system. Retrieved 31 January 2019, from Facebook Code website: https://code.fb.com/core-data/under-the-hood-facebook-s-cold-storage-system/

Bardin, J. (2014, October 6). Why America Is Losing The Online War With ISIS. Retrieved 27 October 2014, from Business Insider website: http://www.businessinsider.com/why-america-is-losing-the-online-war-with-isis-2014-10

Barth, F. (1969). Ethnic groups and boundaries: The social organization of Cultural Difference. Bergen: Universitetsforlaget.

BBC Monitoring. (2014, March 13). Crimeans told: Vote Russia or 'neo-Nazi'. Retrieved from https://www.bbc.com/news/world-europe-26552066

Belluz, J. (2017, May 16). Minnesota is fighting its largest measles outbreak in nearly 30 years. Blame vaccine deniers. - Vox. Retrieved 23 May 2017, from Vox website: https://www.vox.com/2017/5/8/15577316/minnesota-measles-outbreak-explained

Bergh, A. (2015). Seeing is believing; hearing is understanding: Building real trust through virtual tools. Presented at the 20th Iinternational Command and Control Research and Technology Symposium 2015, Annapolis, USA. Retrieved from http://www.dodccrp-test.org/s/071.pdf

Bergh, A. (2018, July). Rebel with a Temporary Cause: The Asymmetrical Access to Distrust, Hipness and Intensity As Resources in Cyber-Conflicts. Conference presentation presented at the XIX ISA World Congress of Sociology, Toronto. Retrieved from https://isaconf.confex.com/isaconf/wc2018/webprogram/Paper104220.html

Bergsaker, A. T., & Bakken, M. L. K. og J. (2018). Smirnoff-is, flaggbot og hai i Stavanger – løgnfabrikker laget Norges mest delte saker i mai. Retrieved 24 February 2019, from Faktisk website: https://www.faktisk.no/artikler/bA/lognfabrikker-laget-norges-mest-delte-saker-i-mai

Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 U.S. Presidential election online discussion. First Monday, 21(11). https://doi.org/10.5210/fm.v21i11.7090

Bessi, A., Scala, A., Rossi, L., Zhang, Q., & Quattrociocchi, W. (2014). The economy of attention in the age of (mis)information. Journal of Trust Management, 1, 12. https://doi.org/10.1186/s40493-014-0012-y

Bialy, B. (2017). New Trends In Social Media. NATO Strategic Communications Centre of Excellence.

Biersack, J., & O'lear, S. (2014). The geopolitics of Russia's annexation of Crimea: Narratives, identity, silences, and energy. Eurasian Geography and Economics, 55(3), 247–269.

Birkbak, A., & Carlsen, H. (2016). The world of Edgerank: Rhetorical justifications of Facebook's News Feed algorithm. Computational Culture, 5(Special Issue on Rhetoric and Computation).

Biswas, S. (2018, August 20). Fighting India's WhatsApp fake news war. Retrieved from https://www.bbc.com/news/world-asia-india-45140158

Bjørnstad, A. L. (2019). Understanding influence in a defense context: A review of relevant research from the field of psychology (FFI-Rapport No. 19/01224). Kjeller, Norway: FFI.

Blomkvist, S. (2002). The User as a Personality-Using Personas as a Tool for Design. KTH-Royal Institute of Technology, Stockholm Www. Nada. Kth. Se/~ Tessy/Blomkvist. Pdf.

Boffey, D. (2018, December 10). Russia 'paved way for Ukraine ship seizures with fake news drive'. The Guardian. Retrieved from https://www.theguardian.com/world/2018/dec/10/russia-paved-way-for-ukraine-ship-seizures-with-fake-news-drive

Bowles, N. (2017, August 30). How ‘Doxxing’ Became a Mainstream Tool in the Culture Wars—The New York Times. Retrieved 6 October 2017, from New York Times website: https://www.nytimes.com/2017/08/30/technology/doxxing-protests.html

Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Bavel, J. J. V. (2017). Emotion shapes the diffusion of moralized content in social networks. Proceedings of the National Academy of Sciences, 201618923. https://doi.org/10.1073/pnas.1618923114

Brandom, R. (2018, February 16). Russia’s troll identities were more sophisticated than anyone thought. Retrieved 7 March 2018, from The Verge website: https://www.theverge.com/2018/2/16/17021684/facebook-twitter-mueller-russia-troll-factory

Braun, K. A., Ellis, R., & Loftus, E. F. (2002). Make my memory: How advertising can change our memories of the past. Psychology and Marketing, 19(1).

Broockman, D. E., & Green, D. P. (2014). Do Online Advertisements Increase Political Candidates’ Name Recognition or Favorability? Evidence from Randomized Field Experiments. Political Behavior, 36(2), 263–289. https://doi.org/10.1007/s11109-013-9239-z

Brooks, R. C. (2017, October 23). How Russians Attempted To Use Instagram To Influence Native Americans [News]. Retrieved 7 March 2018, from BuzzFeed News website: https://www.buzzfeed.com/ryancbrooks/russian-troll-efforts-extended-to-standing-rock

Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. New Media & Society, 14(7), 1164–1180.

Buckels, E. E., Trapnell, P. D., & Paulhus, D. L. (2014). Trolls just want to have fun. Personality and Individual Differences, 67, 97–102.

Cadwalladr, C. (2018, March 17). ‘I created Steve Bannon’s psychological warfare tool’: Meet the data war whistleblower. The Guardian. Retrieved from http://www.theguardian.com/news/2018/mar/17/data-war-whistleblower-christopher-wylie-faceook-nix-bannon-trump

Cadwalladr, C., & Graham-Harrison, E. (2018, March 17). How Cambridge Analytica turned Facebook ‘likes’ into a lucrative political tool. Retrieved 17 March 2018, from The Guardian website: http://www.theguardian.com/technology/2018/mar/17/facebook-cambridge-analytica-kogan-data-algorithm

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. Science, 356(6334), 183–186.

Cellan-Jones, D. L., Rory. (2019, January 21). WhatsApp restricts message-sharing. Retrieved from https://www.bbc.com/news/technology-46945642

Cheaib, A. (2017, October 9). Have You Heard of Click-Farms? Retrieved 26 May 2019, from Ali Cheaib website: https://medium.com/@alicheaib/have-you-heard-of-click-farms-f68eaa465ada

Cheng, J., Bernstein, M., Danescu-Niculescu-Mizil, C., & Leskovec, J. (2017). Anyone Can Become a Troll: Causes of Trolling Behavior in Online Discussions. ArXiv:1702.01119 [Cs, Stat], 1217–1230. https://doi.org/10.1145/2998181.2998213

Christopoulou, D. (2018, March 31). How Norway's Child Welfare Service Is Creating World-Wide Controversy. Retrieved 9 January 2019, from Culture Trip website: https://theculturetrip.com/europe/norway/articles/how-norways-child-welfare-service-is-creating-world-wide-controversy/

Claburn, T. (2017, March 2). Google's troll-destroying AI can't cope with typos. Retrieved 6 October 2017, from https://www.theregister.co.uk/2017/03/02/google_trollspotting_ai_trips_over_typos/

Concordia. (2016). Cambridge Analytica—The Power of Big Data and Psychographics. Retrieved from https://www.youtube.com/watch?v=n8Dd5aVXLCc

Cooper, A., Reimann, R., Cronin, D., & Cooper, A. (2007). About face 3: The essentials of interaction design ([3rd ed.], Completely rev. & updated). Indianapolis, IN: Wiley Pub.

Craker, N., & March, E. (2016). The dark side of Facebook®: The Dark Tetrad, negative social potency, and trolling behaviours. Personality and Individual Differences, 102, 79–84.

Cresci, S., Di Pietro, R., Petrocchi, M., Spognardi, A., & Tesconi, M. (2015). Fame for sale: Efficient detection of fake Twitter followers. Decision Support Systems, 80, 56–71.

Cresci, S., Pietro, R. D., Petrocchi, M., Spognardi, A., & Tesconi, M. (2014). A Fake Follower Story: Improving fake accounts detection on Twitter (No. IIT TR-03/2014). Milan: Consiglio Nazionale delle Ricerche.

Cyber Caliphate: ISIS Plays Offense on the Web. (2015, April 2). Retrieved 27 February 2018, from Recorded Future website: https://www.recordedfuture.com/cyber-caliphate-analysis/

Czarniawska, B. (2004). Narratives in Social Science Research. SAGE.

Dafonte-Gómez, A. (2018). Audiences as Medium: Motivations and Emotions in News Sharing. International Journal of Communication, 12, 20.

Daily Telegraph. (2012, February 9). Leveson Inquiry: The truth behind 'Freddie Starr ate my hamster'. Retrieved 11 March 2019, from https://www.telegraph.co.uk/news/uknews/leveson-inquiry/9072308/Leveson-Inquiry-the-truth-behind-Freddie-Starr-ate-my-hamster.html

Davie, G. (2011, March 17). Framing Theory. Retrieved 16 May 2019, from Mass Communication Theory website: https://masscommtheory.com/theory-overviews/framing-theory/

Deahl, D. (2017, June 12). Three men in Thailand reportedly ran a clickfarm with over 300,000 SIM cards and 400 iPhones—The Verge. Retrieved 26 June 2017, from The Verge website: https://www.theverge.com/2017/6/12/15786402/thai-clickfarm-bust-iphones

DeRosa, J. (2015). Revising the Battle of the Narrative. Journal Article| Jul, 16(12), 51pm.

Devine, C. (2017, October 31). 'Kill them all'—Russian-linked Facebook accounts called for violence. Retrieved 7 March 2018, from CNNMoney website: http://money.cnn.com/2017/10/31/media/russia-facebook-violence/index.html

Dewey, C. (2015, July 16). The fastest-growing 'news' site of 2015 was an obscure content farm for moms [News]. Retrieved 4 February 2019, from Washington Post website: https://www.washingtonpost.com/news/the-intersect/wp/2015/07/16/how-moms-won-the-internet-and-what-that-means-for-the-rest-of-us/

@DFRLab. (2018, March 8). The Russians Who Exposed Russia's Trolls. Retrieved 21 August 2018, from DFRLab website: https://medium.com/dfrlab/the-russians-who-exposed-russias-trolls-72db132e3cd1

Dias, P. (2014). From 'infoxication'to 'infosaturation': A theoretical overview of the cognitive and social effects of digital immersion. Ámbitos. Revista Internacional de Comunicación, n. 24, Año 2014, Primer Trimestre (Primavera).

Dobbin, F. (2009). How Durkheim's Theory of Meaning-making Influenced Organizational Sociology. https://doi.org/10.1093/oxfordhb/9780199535231.003.0009

Donath, J. S. (1999). Identity and deception in the virtual community. Communities in Cyberspace, 1996, 29–59.

Doty, D. H., & Glick, W. H. (1994). Typologies as a unique form of theory building: Toward improved understanding and modeling. Academy of Management Review, 19(2), 230–251.

Dougherty, J. (2014). Everyone lies: The Ukraine conflict and Russia's media transformation. Harvard Kennedy School Shorenstein Center on Media, Politics and Public Policy.

Doward, J., & Gibbs, A. (2017, March 4). Did Cambridge Analytica influence the Brexit vote and the US election? | Politics | The Guardian. Retrieved 27 June 2017, from https://www.theguardian.com/politics/2017/mar/04/nigel-oakes-cambridge-analytica-what-role-brexit-trump

Dunn, A. G., Leask, J., Zhou, X., Mandl, K. D., & Coiera, E. (2015). Associations Between Exposure to and Expression of Negative Opinions About Human Papillomavirus Vaccines on Social Media: An Observational Study. Journal of Medical Internet Research, 17(6). https://doi.org/10.2196/jmir.4343

Dwoskin, E., & Romm, T. (2018, March 19). Facebook's rules for accessing user data lured more than just Cambridge Analytica. Washington Post. Retrieved from https://www.washingtonpost.com/business/economy/facebooks-rules-for-accessing-user-data-lured-more-than-just-cambridge-analytica/2018/03/19/31f6979c-658e-43d6-a71f-afdd8bf1308b_story.html

Editorial Board. (2016, November 25). 'Pizzagate' shows how fake news hurts real people. Retrieved 6 October 2017, from Washington Post website: https://www.washingtonpost.com/opinions/pizzagate-shows-how-fake-news-hurts-real-people/2016/11/25/d9ee0590-b0f9-11e6-840f-e3ebab6bcdd3_story.html

Einstein, M. (2016). Black Ops Advertising: Native Ads, Content Marketing, and the Covert World of the Digital Sell. OR Books.

Etterretningstjenesten. (2019). Fokus 2019: Etterretningstjenestens vurdering av aktuelle sikkerhetsutfordringer. Oslo: Etterretningstjenesten.

Evon, D., & Mikkelson, D. (2017, March 24). Muslim Woman Ignores Dying Victim of London Terror Attack? Retrieved 6 February 2019, from Snopes.com website: https://www.snopes.com/news/2017/03/24/muslim-woman-london-attack/

Faizullaev, A., & Cornut, J. (2017). Narrative practice in international politics and diplomacy: The case of the Crimean crisis. Journal of International Relations and Development, 20(3), 578–604.

Fandos, N., & Roose, K. (2018, August 1). Facebook Identifies an Active Political Influence Campaign Using Fake Accounts. The New York Times. Retrieved from https://www.nytimes.com/2018/07/31/us/politics/facebook-political-campaign-midterms.html

Fiegerman, S. (2018, February 21). In the wake of the Florida shooting, Facebook and Google spread conspiracy theories. Again. Retrieved 7 March 2018, from CNNMoney website: http://money.cnn.com/2018/02/21/technology/facebook-youtube-parkland-conspiracy-theories/index.html

Fisher, A. (2015). Swarmcast: How Jihadist Networks Maintain a Persistent Online Presence. 9(3), 18.

Fisher, M. (2018, December 27). Inside Facebook's Secret Rulebook for Global Political Speech. The New York Times. Retrieved from https://www.nytimes.com/2018/12/27/world/facebook-moderators.html

Fishwick, C. (2014, August 15). How a Polish student's website became an Isis propaganda tool. The Guardian. Retrieved from http://www.theguardian.com/world/2014/aug/15/-sp-polish-man-website-isis-propaganda-tool

Folley, A. (2018, November 13). Ocasio-Cortez shares photo of new 'squad' on Capitol Hill [Text]. Retrieved 31 January 2019, from TheHill website: https://thehill.com/homenews/House/416370-ocasio-cortez-shares-photo-of-her-new-squad-on-capitol-hill

Fort, K., Adda, G., & Cohen, K. B. (2011). Amazon Mechanical Turk: Gold Mine or Coal Mine? Computational Linguistics, 37(2), 413–420. https://doi.org/10.1162/COLI_a_00057

Gabielkov, M., Ramachandran, A., Chaintreau, A., & Legout, A. (2016, June 14). Social Clicks: What and Who Gets Read on Twitter? Presented at the ACM SIGMETRICS / IFIP Performance 2016. Retrieved from https://hal.inria.fr/hal-01281190/document

garson. (2017, July 16). You're Not the Customer; You're the Product. Retrieved 3 February 2019, from Quote Investigator website: https://quoteinvestigator.com/2017/07/16/product/

Geers, K. (2015). Cyber war in perspective: Russian aggression against Ukraine. CCDCOE, NATO Cooperative Cyber Defence Centre of Excellence.

Gibson, J. J. (1977). The Theory of Affordances. In R. Shaw & J. Bransford (Eds.), Perceiving, Acting, and Knowing. Michigan: Lawrence Erlbaum Associates.

Giles, K. (2015). Russia's hybrid Warfare: A success in propaganda. Bundesakademie Für Sicherheitspolitik, Working Paper, (1), 2015.

Gioe, D. V. (2018). Cyber operations and useful fools: The approach of Russian hybrid intelligence. Intelligence and National Security, 1–20.

Goldhaber, M. (2006). The value of openness in an attention economy. First Monday, 11(6).

Goldhaber, M. H. (1997). The attention economy and the net. First Monday, 2(4).

Goodfellow, Ian, Nicolas Papernot, Sandy Huang, Yan Duan, Pieter Abbeel, and Jack Clark. 2017. 'Attacking Machine Learning with Adversarial Examples'. OpenAI Blog. Retrieved 7 March 2018 (https://blog.openai.com/adversarial-example-research/).

Goolsby, R. (2019). Developing a New Approach to Cyber Diplomacy: Addressing Malign Information Maneuvers in Cyberspace. Senior Leadership Roundtable on Military and Defence Aspects of Border Security in South East Europe, 141, 105.

Goolsby, R., & Carley, K. M. (2019). U.S.- Canada Technical Demonstration for NATO Trident Juncture 18—After Action Report (p. 47) [After Action Report]. Arlington, VA: Office of Naval Research.

Graham-Harrison, E., & Cadwalladr, C. (2018, March 17). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. Retrieved 17 March 2018, from The Guardian website: http://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

Grimme, C., Preuss, M., Adam, L., & Trautmann, H. (2017). Social Bots: Human-Like by Means of Human Control? Big Data, 5(4), 279–293. https://doi.org/10.1089/big.2017.0044

Guess, A., Nyhan, B., & Reifler, J. (2018). Selective Exposure to Misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign. European Research Council.

Haciyakupoglu, G., Hui, J. Y., Suguna, V. S., Leong, D., & Rahman, M. F. B. A. (2018). Countering Fake News: A Survey of Recent Global Initiatives.

Hagen, J., & Søgaard, H. A. (2013). Strategisk kommunikasjon som redskap i krisehåndtering (FFI-Rapport No. 2013/03101). Kjeller, Norway: FFI.

Hansen, F. S. (2017). RUSSIAN HYBRID WARFARE: A study of disinformation (No. 2017:06). Copenhagen: DIIS.

Henriksen, A. B. (2017, July 30). Facebook-innlegg om tomme busseter går viralt: - Det ser da virkelig skummelt ut, kan være terrorister med våpen. Retrieved 31 January 2019, from Nettavisen website: http://nettavisen.no/artikkel/3423359933

Henshel, A. (2018, October 25). Twitter loses 9 million monthly active users in Q3 2018, its steepest decline ever. Retrieved 6 February 2019, from VentureBeat website: https://venturebeat.com/2018/10/25/twitter-loses-9-million-monthly-active-users-in-q3-2018-its-steepest-decline-ever/

Hern, A. (2014, October 23). Felicia Day's public details put online after she described Gamergate fears. Retrieved 26 June 2017, from The Guardian website: https://www.theguardian.com/technology/2014/oct/23/felicia-days-public-details-online-gamergate

Hern, A. (2017a, July 12). Extremists driven off Facebook and Twitter targeting smaller firms. The Guardian. Retrieved from http://www.theguardian.com/uk-news/2017/jul/12/extremists-driven-off-facebook-and-twitter-targeting-smaller-firms

Hern, A. (2017b, July 31). Facebook 'dark ads' can swing political opinions, research shows. The Guardian. Retrieved from http://www.theguardian.com/technology/2017/jul/31/facebook-dark-ads-can-swing-opinions-politics-research-shows

Hern, A. (2017c, November 14). How a Russian 'troll soldier' stirred anger after the Westminster attack. The Guardian. Retrieved from https://www.theguardian.com/uk-news/2017/nov/14/how-a-russian-troll-soldier-stirred-anger-after-the-westminster-attack

Hern, A., Duncan, P., & Bengtsson, H. (2017, November 20). Russian 'troll army' tweets cited more than 80 times in UK media. Retrieved 7 March 2018, from The Guardian website: http://www.theguardian.com/media/2017/nov/20/russian-troll-army-tweets-cited-more-than-80-times-in-uk-media

Herrick, D. (2016). The social side of 'cyber power'? Social media and cyber operations. Cyber Conflict (CyCon), 2016 8th International Conference On, 99–111. IEEE.

Herring, S., Job-Sluder, K., Scheckler, R., & Barab, S. (2002). Searching for safety online: Managing" trolling" in a feminist forum. The Information Society, 18(5), 371–384.

Higgins, A. (2016, May 30). Effort to Expose Russia's 'Troll Army' Draws Vicious Retaliation. The New York Times. Retrieved from http://www.nytimes.com/2016/05/31/world/europe/russia-finland-nato-trolls.html

Himler, P. (2013). Content Is King, Distribution Is Queen. Retrieved 11 February 2019, from Forbes website: https://www.forbes.com/sites/peterhimler/2013/07/09/content-is-king-distribution-is-queen/

Hopkins, N. (2017, May 21). Revealed: Facebook's internal rulebook on sex, terrorism and violence. The Guardian. Retrieved from https://www.theguardian.com/news/2017/may/21/revealed-facebook-internal-rulebook-sex-terrorism-violence

Howard, P. N., & Kollanyi, B. (2016). Bots, #Strongerin, and #Brexit: Computational Propaganda During the UK-EU Referendum. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2798311

Hutchings, S., & Szostek, J. (2015). Dominant narratives in Russian political and media discourse during the Ukraine crisis. Ukraine and Russia, 173.

Iasiello, E. J. (2017). Russia's Improved Information Operations: From Georgia to Crimea. Parameters, 47(2).

Irani, L. (2012). Microworking the crowd. Limn, 1(2).

Jacobs, J. (2018, July 19). In Twitter Purge, Top Accounts Lose Millions of Followers. The New York Times. Retrieved from https://www.nytimes.com/2018/07/12/technology/twitter-followers-nyt.html

Jahng, M. R., & Littau, J. (2016). Interacting is believing: Interactivity, social cue, and perceptions of journalistic credibility on twitter. Journalism & Mass Communication Quarterly, 93(1), 38–58.

Jankowicz, N. (2017, March 31). Our biggest mistake in the fight against fake news. Retrieved 27 June 2017, from Washington Post website: https://www.washingtonpost.com/news/democracy-post/wp/2017/03/31/our-biggest-mistake-in-the-fight-against-fake-news/

Joffe, H. (2008). The power of visual material: Persuasion, emotion and identification. Diogenes, 55(1), 84–93.

Jolicoeur, P., & Seaboyer, A. (2018, August). Defence and Security in the Cyber Age; The New Contemporary Operating Environment. Presented at the XIX ISA World Congress of Sociology, Toronto. Retrieved from https://isaconf.confex.com/isaconf/wc2018/webprogram/Paper100660.html

Kalla, J. L., & Broockman, D. E. (2018). The Minimal Persuasive Effects of Campaign Contact in General Elections: Evidence from 49 Field Experiments. American Political Science Review, 112(01), 148–166. https://doi.org/10.1017/S0003055417000363

Kaplan, A. M., & Haenlein, M. (2011). Two hearts in three-quarter time: How to waltz the social media/viral marketing dance. Business Horizons, 54(3), 253–263. https://doi.org/10.1016/j.bushor.2011.01.006

Kaplan, J. T., Gimbel, S. I., & Harris, S. (2016). Neural correlates of maintaining one's political beliefs in the face of counterevidence. Scientific Reports, 6, 39589.

Karlsen, G. H. (2019). Divide and rule: Ten lessons about Russian political influence activities in Europe. Palgrave Communications, 5(1). https://doi.org/10.1057/s41599-019-0227-8

Kastrenakes, J. (2016, May 10). Facebook admits its trending news algorithm needs a lot of human help. Retrieved 6 October 2017, from The Verge website: https://www.theverge.com/2016/5/10/11649296/facebook-explains-human-role-in-trending-topic-selection

Kearns, E., Betus, A., & Lemieux, A. (2017). Why Do Some Terrorist Attacks Receive More Media Attention Than Others? (SSRN Scholarly Paper No. ID 2928138). Retrieved from Social Science Research Network website: https://papers.ssrn.com/abstract=2928138

Keating, D., Schaul, K., & Shapiro, L. (2017, November 1). The Facebook ads Russians targeted at different groups. Retrieved 5 February 2019, from Washington Post website: https://www.washingtonpost.com/graphics/2017/business/russian-ads-facebook-targeting/

Kerwood, P. (2016, January 8). Mr Social on the popularity of Instagram and how to use it. Retrieved 28 August 2018, from Campaign website: https://www.campaignlive.co.uk/article/mr-social-popularity-instagram-use/1374662

Kim, J. W. (2018). Rumor has it: The effects of virality metrics on rumor believability and transmission on Twitter. New Media & Society, 1461444818784945.

Kimball, J. P. (1988). The Stab-in-the-Back Legend and the Vietnam War. Armed Forces & Society, 14(3), 433–458. https://doi.org/10.1177/0095327X8801400306

Kirby, E. J. (2016, December 5). The city getting rich from fake news. BBC News. Retrieved from https://www.bbc.co.uk/news/magazine-38168281

Klang, M. (2016). On The Internet Nobody Can See Your Cape: The ethics of online vigilantism. AoIR Selected Papers of Internet Research, 5.

Kobie, N. (2018, March 20). We were warned about Cambridge Analytica. Why didn't we listen? Retrieved 22 March 2018, from WIRED UK website: http://www.wired.co.uk/article/facebook-cambridge-analytica-data-share-price-privacy

Kozlowska, H., & Kozlowska, H. (2018, February 10). Facebook "likes" are a powerful tool for authoritarian rulers, court petition says. Retrieved 7 March 2018, from Quartz website: https://qz.com/1203637/facebook-likes-are-a-powerful-tool-for-authoritarian-rulers-lawsuit-says/

Krasodomski-Jones, A. (2017). Talking to ourselves: Political debate online and the echo chamber effect. Retrieved from https://demosuk.wpengine.com/wp-content/uploads/2017/02/Echo-Chambers-final-version.pdf

Krauss, S. E. (2005). Research paradigms and meaning making: A primer. The Qualitative Report, 10(4), 758–770.

Lapidot-Lefler, N., & Barak, A. (2012). Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition. Computers in Human Behavior, 28(2), 434–443.

Larson, E. V., Darilek, R. E., Gibran, D., Nichiporuk, B., Richardson, A., Schwartz, L. H., & Thurston, C. Q. (2009). Foundations of effective influence operations: A framework for enhancing army capabilities. RAND ARROYO CENTER SANTA MONICA CA.

Layton, P. (2018, April 25). Duelling Algorithms: Using Artificial Intelligence in Warfighting. Retrieved 10 August 2018, from OTH website:

https://othjournal.com/2018/04/25/duelling-algorithms-using-artificial-intelligence-in-warfighting/

Lee, D. (2018, March 22). YouTube gun ban drives bloggers to PornHub. BBC News. Retrieved from http://www.bbc.com/news/technology-43500714

Lee, H., & Oh, H. J. (2017). Normative mechanism of rumor dissemination on Twitter. Cyberpsychology, Behavior, and Social Networking, 20(3), 164–171.

Leppänen, S., Kytölä, S., Jousmäki, H., Peuronen, S., & Westinen, E. (2013). Entextualization and resemiotization as resources for (dis)identification in social media. Retrieved from https://www.tilburguniversity.edu/upload/a3d5524e-4413-4772-9f96-9fe0ee714c6f_TPCS_57_Leppanen-etal.pdf

Lessig, L. (2008). Remix: Making art and commerce thrive in the hybrid economy. Penguin.

Levin, S. (2017, October 2). Facebook and Google promote politicized fake news about Las Vegas shooter | US news | The Guardian. Retrieved 6 October 2017, from https://www.theguardian.com/us-news/2017/oct/02/las-vegas-shooting-facebook-google-fake-news-shooter

Levine, J. (2018, June 4). Roseanne Twitter Followers Jump More Than 35 Percent After Show Cancellation. Retrieved 4 February 2019, from MSN website: https://www.msn.com/en-us/tv/news/roseanne-twitter-followers-jump-more-than-35-percent-after-show-cancellation/ar-AAye1cK

Lewis, P. (2018a, February 2). 'Fiction is outperforming reality': How YouTube's algorithm distorts truth. The Guardian. Retrieved from http://www.theguardian.com/technology/2018/feb/02/how-youtubes-algorithm-distorts-truth

Lewis, P. (2018b, March 20). 'Utterly horrifying': Ex-Facebook insider says covert data harvesting was routine. Retrieved 20 March 2018, from The Guardian website: http://www.theguardian.com/news/2018/mar/20/facebook-data-cambridge-analytica-sandy-parakilas

Lister, T., Sciutto, J., & Ilyushina, M. (2017, October 17). Putin's 'chef,' the man behind the troll factory. Retrieved 7 March 2018, from CNN website: https://www.cnn.com/2017/10/17/politics/russian-oligarch-putin-chef-troll-factory/index.html

Lucas, E., & Nimmo, B. (2015). Information Warfare: What Is It and How to Win It? (CEPA INFOWAR PAPER No. No. 1). Retrieved from http://cepa.org/sites/default/files/Infowar%20Report.pdf

Mandell, A. (2018, May 30). Inside Roseanne Barr's history of offensive tweets. Retrieved 4 February 2019, from USA Today website: http://www.usatoday.com/story/life/people/2018/05/29/roseanne-barr-offensive-tweeting-actually-nothing-new/652292002/

Mann, T. (2017, June 4). Four blasts heard near London Bridge [News]. Retrieved 6 February 2019, from Metro website: https://metro.co.uk/2017/06/04/three-explosions-heard-near-london-bridge-6682987/

Manthorpe, R. (2018, March 20). The UK's left is scrambling to adapt to Facebook's algorithm change. Retrieved 22 March 2018, from WIRED UK website: http://www.wired.co.uk/article/facebook-algorithm-changes-engagement-labour

Markham, A. N. (2003). Metaphors Reflecting and Shaping the Reality of the Internet: Tool, Place, Way of Being. Retrieved from https://pure.au.dk/portal/files/69632404/MarkhamTPW.pdf

Marty, R. (2018, February 12). How Google Maps Leads Women Seeking Abortions Astray. Retrieved 7 March 2018, from Gizmodo website: https://gizmodo.com/how-google-maps-leads-women-seeking-abortions-astray-1822882758

Massanari, A. L. (2010). Designing for imaginary friends: Information architecture, personas and the politics of user-centered design. New Media & Society, 12(3), 401–416.

Matejic, N. (2016). Content Wars: Daesh's sophisticated use of communications. Retrieved 25 November 2016, from NATO Review website: http://www.nato.int/docu/review/2016/Also-in-2016/wars-media-daesh-communications-solis/EN/index.htm

Matthias Niessner. (2016). Face2Face: Real-time Face Capture and Reenactment of RGB Videos (CVPR 2016 Oral). Retrieved from https://www.youtube.com/watch?time_continue=182&v=ohmajJTcpNk

Mazzetti, M., & Gordon, M. R. (2015, June 12). ISIS Is Winning the Social Media War, U.S. Concludes. The New York Times. Retrieved from http://www.nytimes.com/2015/06/13/world/middleeast/isis-is-winning-message-war-us-concludes.html

McHugh, L. (2015). The Role of Social Media in the ISIL-West Crisis: A Technoethical Analysis of Twitter. Retrieved from https://www.ruor.uottawa.ca/handle/10393/34898

McKay, T. (2017, November 5). Once Again, Google Promoted Disinformation and Propaganda After a Mass Shooting [Updated]. Retrieved 7 March 2018, from Gizmodo website: https://gizmodo.com/once-again-google-promoted-disinformation-and-propagan-1820166979/amp

McKew, M. (2018, February 16). Did Russia Affect the 2016 Election? It's Now Undeniable | WIRED. Retrieved 27 February 2018, from Wired website: https://www.wired.com/story/did-russia-affect-the-2016-election-its-now-undeniable/

Mellor, C. (2013, September 24). Facebook Frankenphoto morgue will store your cold, dead selfies FOREVER. Retrieved 25 January 2019, from https://www.theregister.co.uk/2013/09/24/facebook_on_the_rue_morgue/

Merrill, S., & Åkerlund, M. (2018). Standing Up for Sweden? The Racist Discourses, Architectures and Affordances of an Anti-Immigration Facebook Group. Journal of

Computer-Mediated Communication, 23(6), 332–353. https://doi.org/10.1093/jcmc/zmy018

Mezzofiore, G. (2018, February 16). Russian bots promote guns after Florida shooting. Retrieved 7 March 2018, from CNN website: https://www.cnn.com/2018/02/16/us/russian-bots-florida-shooting-intl/index.html

Molden, D. T. (2015). How Do You Catch a Cloud and Pin it Down? The struggle to define and identify the GamerGate 'movement'. Graduate School Proceedings, 7, 41–56. Aichi Shukutoku University.

Mørch, W.-T. (2014, December 30). Han kobler det norske barnevernet til terror og overgrep. Nordlys. Retrieved from https://nordnorskdebatt.no/article/han-kobler-norske-barnevernet

Morgan, J. (2017, November 6). How to Fool Americans on Twitter. Retrieved 30 January 2019, from Data for Democracy website: https://medium.com/data-for-democracy/how-to-fool-americans-on-twitter-2a1da10724a2

Morris, M. R., Counts, S., Roseway, A., Hoff, A., & Schwarz, J. (2012). Tweeting is believing?: Understanding microblog credibility perceptions. Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work, 441–450. ACM.

Napoli, P. M. (2014). On automation in media industries: Integrating algorithmic media production into media industries scholarship. Media Industries Journal, 1(1).

Nicander, L. (2001). Information Operations–A Swedish View. Journal of Information Warfare, 1(1), 16–25.

Oh, O., Agrawal, M., & Rao, H. R. (2013). Community intelligence and social media services: A rumor theoretic analysis of tweets during social crises. MIS Quarterly, 37(2), 407–426.

Oremus, W. (2016, December 14). An interview with Charles Delavan, the IT guy whose "typo" led to the Podesta email hack. Retrieved 23 July 2018, from Slate website: http://www.slate.com/articles/technology/future_tense/2016/12/an_interview_with_charles_delavan_the_it_guy_whose_typo_led_to_the_podesta.html?via=gdpr-consent#lf_comment=621987277

O'Sullivan, D. (2017, October 6). Exclusive: Russian-linked group sold merchandise online. Retrieved 7 March 2018, from CNNMoney website: http://money.cnn.com/2017/10/06/media/blacktivist-merchandise-facebook-russia/index.html

O'Sullivan, D., Devine, C., & Griffin, D. (2018, March 26). Obama official: We could have stopped Russian trolls. Retrieved 27 March 2018, from CNN website: https://www.cnn.com/2018/03/26/politics/brett-bruen-russian-meddling-election/index.html

Oxford Dictionaries. (2018). Definition of useful idiot in English by Oxford Dictionaries. Retrieved 12 June 2018, from Oxford Dictionaries | English website: https://en.oxforddictionaries.com/definition/useful_idiot

Papacharissi, Z. (2015). Toward new journalism (s) affective news, hybridity, and liminal spaces. Journalism Studies, 16(1), 27–40.

Parakilas, S. (2018, March 20). I worked at Facebook. I know how Cambridge Analytica could have happened. Washington Post. Retrieved from https://www.washingtonpost.com/opinions/i-worked-at-facebook-i-know-how-cambridge-analytica-could-have-happened/2018/03/20/edc7ef8a-2bc4-11e8-8ad6-fbc50284fce8_story.html

Parkin, S. (2014a). Gamergate: A scandal erupts in the video-game community. The New Yorker, 17.

Parkin, S. (2014b, September 9). Zoe Quinn's Depression Quest. The New Yorker. Retrieved from https://www.newyorker.com/tech/elements/zoe-quinns-depression-quest

Parlapiano, A. (2018, February 16). The Propaganda Tools Used by Russians to Influence the 2016 Election. The New York Times. Retrieved from https://www.nytimes.com/interactive/2018/02/16/us/politics/russia-propaganda-election-2016.html, https://www.nytimes.com/interactive/2018/02/16/us/politics/russia-propaganda-election-2016.html

Paul, C., & Matthews, M. (2016). The Russian 'Firehose of Falsehood' Propaganda Model: Why It Might Work and Options to Counter It. https://doi.org/10.7249/PE198

Paviour, B. (2017, October 31). What a Facebook experiment did to news in Cambodia. BBC News. Retrieved from http://www.bbc.com/news/world-asia-41801071

Pavleska, T., Školkay, A., Zankova, B., Ribeiro, N., & Bechmann, A. (2018). Performance analysis of fact-checking organizations and initiatives in Europe: A critical overview of online platforms fighting fake news.

Pennycook, G., Cannon, T., & Rand, D. G. (2017). Implausibility and Illusory Truth: Prior Exposure Increases Perceived Accuracy of Fake News but Has No Effect on Entirely Implausible Statements (SSRN Scholarly Paper No. ID 2958246). Retrieved from Social Science Research Network website: https://papers.ssrn.com/abstract=2958246

Pennycook, G., & Rand, D. G. (2017). The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Stories Increases Perceived Accuracy of Stories Without Warnings (SSRN Scholarly Paper No. ID 3035384). Retrieved from Social Science Research Network website: https://papers.ssrn.com/abstract=3035384

Penzenstadler, N., Heath, B., & Guynn, J. (2018, May 13). What we found in Facebook ads by Russians accused of election meddling. USA Today. Retrieved from http://www.usatoday.com/story/news/2018/05/11/what-we-found-facebook-ads-russians-accused-election-meddling/602319002/

Permanent Select Committee on Intelligence. (2018). Social Media Advertisements. Retrieved 6 February 2019, from https://intelligence.house.gov/social-media-content/social-media-advertisements.htm

Perry, B. (2015). Non-linear warfare in Ukraine: The critical role of information operations and special operations. Small Wars Journal, 14, 1–30.

Pilkington, E., & Michel, A. (2012, February 17). Obama, Facebook and the power of friendship: The 2012 data election. Retrieved 22 March 2018, from The Guardian website: http://www.theguardian.com/world/2012/feb/17/obama-digital-data-machine-facebook-election

PM, T. W. O. 10/25/14 at 12:32. (2014, October 25). Is GamerGate About Media Ethics or Harassing Women? Harassment, the Data Shows. Retrieved 20 March 2018, from Newsweek website: http://www.newsweek.com/gamergate-about-media-ethics-or-harassing-women-harassment-data-show-279736

Poulsen, K., & Ackerman, S. (2018, March 22). 'Lone DNC Hacker' Guccifer 2.0 Slipped Up and Revealed He Was a Russian Intelligence Officer. The Daily Beast. Retrieved from https://www.thedailybeast.com/exclusive-lone-dnc-hacker-guccifer-20-slipped-up-and-revealed-he-was-a-russian-intelligence-officer

Prior, N. (2018). Popular Music, Digital Technology and Society. SAGE.

Pruitt, J., & Adlin, T. (2006). The Persona Lifecycle: Keeping People in Mind Throughout Product Design (1 edition). Amsterdam ; Boston: Morgan Kaufmann.

Quach, K. (2018, September 11). The eyes don't have it! AI's 'deep-fake' vids surge ahead in realism [IT News]. Retrieved 12 September 2018, from The Register website: https://www.theregister.co.uk/2018/09/11/ai_fake_videos/

Rácz, A. (2015). Russia's Hybrid War in Ukraine: Breaking the Enemy's Ability to Resist (FIIA Report No. 43; p. 104). Helsinki.

Reinstein, J. (2018, January 12). 'Tweetdecking' Is Taking Over Twitter. Here's Everything You Need To Know. Retrieved 12 March 2018, from BuzzFeed website: https://www.buzzfeed.com/juliareinstein/exclusive-networks-of-teens-are-making-thousands-of-dollars

Richardson, J. E., & Wodak, R. (2009). The Impact of Visual Racism: Visual Arguments in Political Leaflets of Austrian and British Far-right Parties. Controversia, 6(2).

Richter, M. L. (2017). The Kremlin's Platform for 'Useful Idiots' in the West: An Overview of RT's Editorial Strategy and Evidence of Impact. Praha: European Values.

Robertson, A. (2014, August 27). Trolls drive Anita Sarkeesian out of her house to prove misogyny doesn't exist. Retrieved 11 July 2018, from The Verge website: https://www.theverge.com/2014/8/27/6075179/anita-sarkeesian-says-she-was-driven-out-of-house-by-threats

Rourke, D. (2015). Algorithmic Narratives and Synthetic Subjects. Daniel Rourke. Presented at the Theorizing the Web Conference. Retrieved from https://machinemachine.net/portfolio/paper-at-theorizing-the-web-synthetic-subjects/

RT.com. (2015). Roseanne Barr talks about MK ULTRA mind control in Hollywood. Retrieved from https://www.youtube.com/watch?v=n6wxEa8RRW4

Ruddick, G. (2017, August 18). Experts sound alarm over news websites' fake news twins. Retrieved 7 March 2018, from The Guardian website: http://www.theguardian.com/technology/2017/aug/18/experts-sound-alarm-over-news-websites-fake-news-twins

Safi, M. (2018, March 14). Sri Lanka accuses Facebook over hate speech after deadly riots. Retrieved 25 March 2018, from The Guardian website: http://www.theguardian.com/world/2018/mar/14/facebook-accused-by-sri-lanka-of-failing-to-control-hate-speech

Safi, M., & Perera, A. (2018, March 7). Sri Lanka blocks social media as deadly violence continues. Retrieved 9 March 2018, from The Guardian website: http://www.theguardian.com/world/2018/mar/07/sri-lanka-blocks-social-media-as-deadly-violence-continues-buddhist-temple-anti-muslim-riots-kandy

Sample, I. (2018, February 20). Bad News: The game researchers hope will 'vaccinate' public against fake news. Retrieved 7 March 2018, from The Guardian website: http://www.theguardian.com/technology/2018/feb/20/bad-news-the-game-researchers-hope-will-vaccinate-public-against-fake-news

Santa Maria, S. D. (2013). Improving Influence Operations by Defining Influence and Influence Operations. https://doi.org/10.21236/ADA606282

Satter, R., & Vlasov, D. (2017, May 11). Ukraine soldiers bombarded by 'pinpoint propaganda' texts. Retrieved 22 May 2017, from AP News website: https://apnews.com/9a564a5f64e847d1a50938035ea64b8f/Sinister-text-messages-reveal-high-tech-front-in-Ukraine-war

Seddon, M. (2014). Documents Show How Russia's Troll Army Hit America. Retrieved 27 March 2018, from BuzzFeed website: https://www.buzzfeed.com/maxseddon/documents-show-how-russias-troll-army-hit-america

Seo, H. (2014). Visual propaganda in the age of social media: An empirical analysis of Twitter images during the 2012 Israeli–Hamas conflict. Visual Communication Quarterly, 21(3), 150–161.

Serrano-Puche, J. (2016). Internet and emotions: New trends in an emerging field of research.

Shane, S., & Mazzetti, M. (2018, February 16). Inside a 3-Year Russian Campaign to Influence U.S. Voters. The New York Times. Retrieved from https://www.nytimes.com/2018/02/16/us/politics/russia-mueller-election.html

Shao, C., Ciampaglia, G. L., Varol, O., Flammini, A., & Menczer, F. (2017). The spread of fake news by social bots. ArXiv Preprint ArXiv:1707.07592.

Shapiro, Leslie. (2017, November 1). Anatomy of a Russian Facebook ad. Retrieved 7 March 2018, from Washington Post website: https://www.washingtonpost.com/graphics/2017/business/russian-ads-facebook-anatomy/

Shin, D.-H., An, H., & Kim, J. H. (2016). How the Second Screens Change the Way People Interact and Learn: The Effects of Second Screen Use on Information Processing. Interactive Learning Environments, 24(8), 2058–2079. https://doi.org/10.1080/10494820.2015.1076851

Siddiqui, F., & Svrluga, S. (2016, December 5). N.C. man told police he went to D.C. pizzeria with gun to investigate conspiracy theory—The Washington Post. Retrieved 13 February 2019, from Washington Post website: https://www.washingtonpost.com/news/local/wp/2016/12/04/d-c-police-respond-to-report-of-a-man-with-a-gun-at-comet-ping-pong-restaurant/?utm_term=.b4f47d1aaf80

Simon, H. (1971). Designing Organizations for an Information-Rich World. In M. Greenberger (Ed.), Computers, communications, and the public interest (pp. 37–72). Baltimore: Johns Hopkins Press.

Sindelar, D. (2014, August 12). The Kremlin's Troll Army. Retrieved 21 August 2018, from The Atlantic website: https://www.theatlantic.com/international/archive/2014/08/the-kremlins-troll-army/375932/

Sinders, C. (2015, July 17). That Time the Internet Sent a SWAT Team to My Mom's House. Retrieved 26 June 2017, from Narratively website: http://narrative.ly/that-time-the-internet-sent-a-swat-team-to-my-moms-house/

Sitawarin, C., Bhagoji, A. N., Mosenia, A., Chiang, M., & Mittal, P. (2018). DARTS: Deceiving Autonomous Cars with Toxic Signs. ArXiv:1802.06430 [Cs]. Retrieved from http://arxiv.org/abs/1802.06430

Sloane, G. (2016, June 7). What are dark posts? Retrieved 6 October 2017, from Digiday website: https://digiday.com/media/what-are-dark-posts/

Smith, K. B. (2002). Typologies, Taxonomies, and the Benefits of Policy Classification. Policy Studies Journal, 30(3), 379–395. https://doi.org/10.1111/j.1541-0072.2002.tb02153.x

Spillman, L. (2002). Introduction: Culture and cultural sociology. Cultural Sociology, 1–16.

Stella, M., Ferrara, E., & De Domenico, M. (2018). Bots sustain and inflate striking opposition in online social systems. ArXiv Preprint ArXiv:1802.07292.

Stelter, B. (2008, March 27). Finding Political News Online, the Young Pass It On. The New York Times. Retrieved from http://www.nytimes.com/2008/03/27/us/politics/27voters.html

Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and Information Diffusion in Social Media—Sentiment of Microblogs and Sharing Behavior. Journal of Management Information Systems, 29(4), 217–248. https://doi.org/10.2753/MIS0742-1222290408

Suler, J. (2004). The online disinhibition effect. Cyberpsychology & Behavior, 7(3), 321–326.

Suler, J. (2005). The online disinhibition effect (2). International Journal of Applied Psychoanalytic Studies, 2(2), 184–188.

Svensson, J. (2013). Power, Identity, and Feelings in Digital Late Modernity: The Rationality of Reflexive Emotion Displays Online. In Internet and Emotions (pp. 31–46). Routledge.

Svetoka, S. (2016). Social Media as a Tool of Hybrid Warfare. NATO Strategic Communications Centre of Excellence.

Tangen, H. B. (2017). Du vil ikke tro hvordan titlene i sportsjournalistikk varierer fra nettutgaven til papiravisa-En studie av sportsjournalistikkens titler i nett-og papiraviser (B.S. thesis). HiOA.

Tatham, S. (2015). Target Audience Analysis. Retrieved from StratCom Centre of Excellence website: http://www.jwc.nato.int/images/stories/threeswords/TAA.pdf

Techopedia. (2019). What is a Platform? - Definition from Techopedia. Retrieved 12 February 2019, from Techopedia.com website: https://www.techopedia.com/definition/3411/platform

Temperton, J. (2018, March 21). This is the smoking gun at the centre of the Facebook and Cambridge Analytica story [News]. Retrieved 22 March 2018, from WIRED UK website: http://www.wired.co.uk/article/facebook-cambridge-analytica-mark-zuckerberg-mission-data-privacy

Terranova, T. (2000). Free Labor: Producing Culture for the Digital Economy. Social Text, 18(2), 33–58.

Terranova, T. (2012). Attention, economy and the brain. Culture Machine, 13.

Timberg, C. (2017, February 5). As a conservative Twitter user sleeps, his account is hard at work. Retrieved 6 February 2017, from Washington Post website: https://www.washingtonpost.com/business/economy/as-a-conservative-twitter-user-sleeps-his-account-is-hard-at-work/2017/02/05/18d5a532-df31-11e6-918c-99ede3c8cafa_story.html

Timberg, C., & Dwoskin, E. (2018, July 6). Twitter is sweeping out fake accounts like never before, putting user growth at risk. Retrieved 6 February 2019, from Washington Post website: https://www.washingtonpost.com/technology/2018/07/06/twitter-is-sweeping-out-fake-accounts-like-never-before-putting-user-growth-risk/

Timberg, C., Dwoskin, E., Entous, A., & Demirjian, K. (2017, November 1). Russian ads, now publicly released, show sophistication of influence campaign. Washington Post. Retrieved from https://www.washingtonpost.com/business/technology/russian-ads-now-publicly-released-show-sophistication-of-influence-campaign/2017/11/01/d26aead2-bf1b-11e7-8444-a0d4f04b89eb_story.html

Timberg, C., & Romm, T. (2018, March 1). These provocative images show Russian trolls sought to inflame debate over climate change, fracking and Dakota pipeline. Washington Post. Retrieved from https://www.washingtonpost.com/news/the-switch/wp/2018/03/01/congress-russians-trolls-sought-to-inflame-u-s-debate-on-climate-change-fracking-and-dakota-pipeline/

TKOBeauty. (2019, January 27). Image of Ocasio-Cortez and new congresswomen. Retrieved 31 January 2019, from https://twitter.com/TKOBeauty/status/1089697386115551232

Treré, E. (2016). The Dark Side of Digital Politics: Understanding the Algorithmic Manufacturing of Consent and the Hindering of Online Dissidence. IDS Bulletin, 47(1). Retrieved from http://bulletin.ids.ac.uk/idsbo/article/view/41/html

Treverton, G. F. (2017). Influence Operations and the Intelligence Policy Challenges. Retrieved from Center for Asymmetric Threat Studies website: https://www.fhs.se/download/18.1ee9003b162cad2caa5351cf/1524483543405/Influence%20Operations%20and%20the%20Intelligence%20Policy%20Challenges.pdf

Treverton, G. F., & Miles, R. (2014). Social media and intelligence. Center for Asymmetric Threat Studies (CATS), Swedish National Defence College.

Tsikerdekis, M. (2012). The choice of complete anonymity versus pseudonymity for aggression online. Int J Hum-Comput Int, 2(8), 35–57.

Tufekci, Z. (2013). "Not this one" social movements, the attention economy, and microcelebrity networked activism. American Behavioral Scientist, 57(7), 848–870.

Tunnicliffe, I., & Tatham, S. (2017). Social Media: The Vital Ground, Can We Hold It?

Tynan, D. (2016, August 24). How Facebook powers money machines for obscure political 'news' sites. The Guardian. Retrieved from https://www.theguardian.com/technology/2016/aug/24/facebook-clickbait-political-news-sites-us-election-trump

Udris, R. (2014). Cyberbullying among high school students in Japan: Development and validation of the Online Disinhibition Scale. Computers in Human Behavior, 41, 253–261.

Usher, N. (2010). Why spreadable doesn't equal viral: A conversation with Henry Jenkins. Nieman Journalism Lab, November, 23.

Valenti, J. (2015, August 29). Anita Sarkeesian interview: 'The word "troll" feels too childish. This is abuse'. The Guardian. Retrieved from http://www.theguardian.com/technology/2015/aug/29/anita-sarkeesian-gamergate-interview-jessica-valenti

Van Cauwenberge, A., Schaap, G., & Van Roy, R. (2014). "TV no longer commands our full attention": Effects of second-screen viewing and task relevance on cognitive load and learning from news. Computers in Human Behavior, 38, 100–109.

Wakefield, A., Murch, S., Anthony, A., Linnell, J., Casson, D., Malik, M., … Walker-Smith, J. (1998). RETRACTED: Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and

pervasive developmental disorder in children. The Lancet, 351(9103), 637–641. https://doi.org/10.1016/S0140-6736(97)11096-0

Walker, S. (2015, April 2). Salutin' Putin: Inside a Russian troll house. The Guardian. Retrieved from http://www.theguardian.com/world/2015/apr/02/putin-kremlin-inside-russian-troll-house

Wallaroo Media. (2014, September 22). Facebook News Feed Algorithm History | 2017 Update. Retrieved 11 January 2018, from Wallaroo Media website: https://wallaroomedia.com/facebook-newsfeed-algorithm-change-history/

Walter, E. (2012, August 28). The Rise Of Visual Social Media [News]. Retrieved 15 August 2018, from Fast Company website: https://www.fastcompany.com/3000794/rise-visual-social-media

Weedon, J., Nuland, W., & Stamos, A. (2017). Information Operations and Facebook (p. 13).

Wen, E., Cao, J., Shen, J., & Liu, X. (2018). Fraus: Launching Cost-efficient and Scalable Mobile Click Fraud Has Never Been So Easy. 2018 IEEE Conference on Communications and Network Security (CNS), 1–9. https://doi.org/10.1109/CNS.2018.8433126

Wendling, M. (2018, January 22). The (almost) complete history of 'fake news'. BBC News. Retrieved from http://www.bbc.com/news/blogs-trending-42724320

Whewell, T. (2018a, April 14). Norway's Barnevernet: They took our four children… then the baby—BBC News. Retrieved 29 August 2018, from BBC News website: https://www.bbc.co.uk/news/magazine-36026458

Whewell, T. (2018b, August 3). Norway's hidden scandal. Retrieved 29 August 2018, from BBC News website: https://www.bbc.co.uk/news/resources/idt-sh/norways_hidden_scandal

Williams, H. T. P., McMurray, J. R., Kurz, T., & Hugo Lambert, F. (2015). Network analysis reveals open forums and echo chambers in social media discussions of climate change. Global Environmental Change, 32, 126–138. https://doi.org/10.1016/j.gloenvcha.2015.03.006

Williams, M., & Burnap, P. (2018). Antisemitic Content on Twitter. Retrieved from Community Security Trust website: https://cst.org.uk/public/data/file/4/2/Antisemitic%20Content%20on%20Twitter.pdf

Worley, M., & Copsey, N. (2016). White Youth: The Far Right, Punk and British Youth Culture, 1977-87. JOMEC Journal, 0(9), 27. https://doi.org/10.18573/j.2016.10041

About FFI

The Norwegian Defence Research Establishment (FFI) was founded 11th of April 1946. It is organised as an administrative agency subordinate to the Ministry of Defence.

FFI's MISSION

FFI is the prime institution responsible for defence related research in Norway. Its principal mission is to carry out research and development to meet the requirements of the Armed Forces. FFI has the role of chief adviser to the political and military leadership. In particular, the institute shall focus on aspects of the development in science and technology that can influence our security policy or defence planning.

FFI's VISION

FFI turns knowledge and ideas into an efficient defence.

FFI's CHARACTERISTICS

Creative, daring, broad-minded and responsible.

Om FFI

Forsvarets forskningsinstitutt ble etablert 11. april 1946. Instituttet er organisert som et forvaltningsorgan med særskilte fullmakter underlagt Forsvarsdepartementet.

FFIs FORMÅL

Forsvarets forskningsinstitutt er Forsvarets sentrale forskningsinstitusjon og har som formål å drive forskning og utvikling for Forsvarets behov. Videre er FFI rådgiver overfor Forsvarets strategiske ledelse. Spesielt skal instituttet følge opp trekk ved vitenskapelig og militærteknisk utvikling som kan påvirke forutsetningene for sikkerhetspolitikken eller forsvarsplanleggingen.

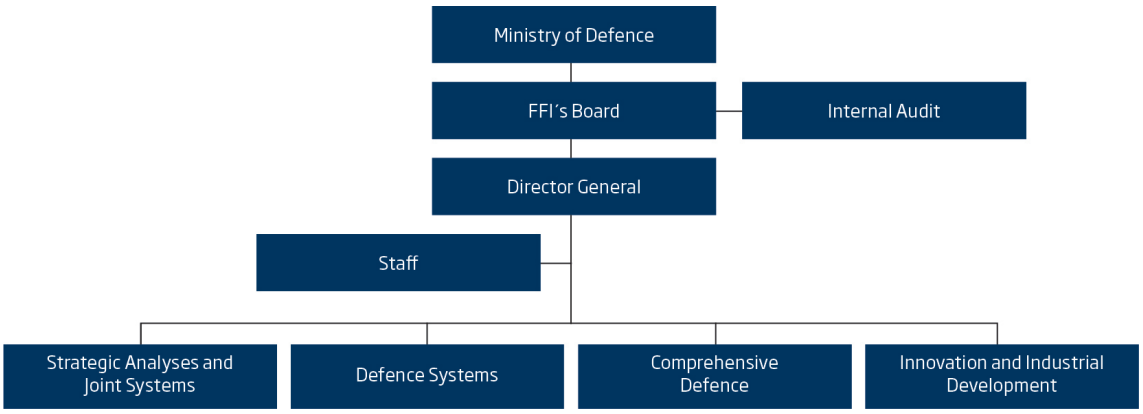
FFIs VISJON

FFI gjør kunnskap og ideer til et effektivt forsvar.

FFIs VERDIER

Skapende, drivende, vidsynt og ansvarlig.

FFI's organisation



Forsvarets forskningsinstitutt
Postboks 25
2027 Kjeller

Besøksadresse:
Instituttveien 20
2007 Kjeller

Telefon: 63 80 70 00
Telefaks: 63 80 71 15
Epost: ffi@ffi.no

Norwegian Defence Research Establishment (FFI)
P.O. Box 25
NO-2027 Kjeller

Office address:
Instituttveien 20
N-2007 Kjeller

Telephone: +47 63 80 70 00
Telefax: +47 63 80 71 15
Email: ffi@ffi.no

FFI Forsvarets forskningsinstitutt
Norwegian Defence Research Establishment