

# The politics of social media manipulation: A view from the Netherlands

Richard Rogers and Sabine Niederer<sup>1</sup>

## Summary

Disinformation and fake news are contemporary phenomena with rich histories. Disinformation, or the wilful introduction of false information for the purposes of causing harm, recalls infamous foreign interference operations in national media systems, such as the Russian campaign 'Operation Infektion' that in the early 1980s effectively publicly linked the HIV virus with a supposed, secret US bioweapons lab. Outcries over fake news, or dubious stories that have the trappings of news, have occurred repeatedly with the introduction of new media technologies that disrupt the publication, distribution and consumption of news -- from the so-called rumour-mongering broadsheet centuries ago to the blogosphere more recently. Designating a news organization as fake, or calling it *der Lügenpresse*, however, has a darker history, associated with authoritarian regimes or populist bombast diminishing the reputation of 'elite media' and the value of inconvenient truths more generally.

These days social media platforms have been implicated in both the practice of disinformation as well as the rise of the two varieties of fake news. As discussed in the theoretical and empirical scholarship to date, social media have enabled the penetration of foreign disinformation operations, the widespread publication and spread of dubious content as well as extreme commentators with considerable followings attacking mainstream media as fake.

Worldwide, disinformation and fake news are increasingly under study together, but the argument could be made to separate them. Indeed, in the Netherlands evidence of foreign disinformation campaigning is scant; cases of domestic actors employing the 'Russian playbook' of disinformation tactics are also seldom documented. Unlike in the case of the US, to which much of the scholarship is dedicated, in the Netherlands one has not witnessed the rise of imposter news organisations or advocacy groups purporting to represent social groups or causes. Indeed, when employing narrow definitions of disinformation and fake news, there is hardly any to be found in the Netherlands.

But definitions of fake and its next-of-kin 'junk news' often extend to clickbait, hyperpartisan sources as well as artificially amplified social media content and accounts. As a case in point, when *Buzzfeed News* famously reported in 2016 that 'fake news' was outperforming mainstream news on Facebook, included in its definition were clickbait and hyperpartisan sources. Expanding the definition in such a manner would have consequences in that the Netherlands has all of them in relative abundance.

Initial studies have found that the Dutch are great consumers of clickbait and 'pulp' content; there is a well engaged-with set of tendentious and highly partisan news-like organisations especially on the right of the political spectrum, and the artificial amplification of social media accounts, including those of certain politicians and musicians, has been well documented. Their sway varies. Clickbait is said to be consumed more often than mainstream news, though there is also more of it. News furnished by commercial and public broadcasting are still referenced, liked or shared in greater

---

<sup>1</sup> Media Studies, University of Amsterdam and Digital Media and Creative Industries, Amsterdam University of Applied Sciences. Version dated 9 April 2019.

quantities than tendentious and hyperpartisan sources, though the latter have been present in the most engaged-with lists of sources around election issues. Artificial amplification both burnishes one's image but also has led to mini-scandals when fake followers are revealed through new online detection tools and news reporting.

Whether any of them is particularly persuasive is a question increasingly posed. The first wave of scholarship on the production and spread of disinformation has yielded to the next wave on its effects. Given people's hardened attitudes the question concerns whether the influence of disinformation and fake news is ever more than 'minimal'.

In that regard, the rise of extreme content (including extreme clickbait), circulated on social media platforms, is one source of continuing consternation and measurement, leading to calls for platform regulation. Another is the mainstreaming of doubt and trust in public institutions and media, concomitant with the rise of both 'alternative facts' and 'alternative fact infrastructures'. The post-truth condition, as it is termed, is discussed as both first-order 'fact fights' as well as second-order competitions between 'sectarian knowledge' regimes and competing media ecologies. Is the authority of mainstream news and knowledge institutions declining for increasing segments of society that consume the alternatives? One finding often related is that older consumers are 'available audiences' for fringe media, and are relatively 'heavy users'.

The consuming and sharing of fake news have been the subject of media literacy initiatives, including quizzes, serious games and public service campaigning. Through heightened awareness, especially around the time of elections, the impact on consumers of any disinformation and dubious content may be mitigated and the institutions made resilient, it has been argued. Voluntary and professional fact-checking are also discussed in this regard, as are (automated) content flagging, together with the need for human review. The question regularly posed is whether the sheer amount of platform junk will overwhelm the capacity to review it, together with the related issue of who will review the reviewers.

Finally, there is widespread scholarly concern about the restrictions to public access of social media data, otherwise known as the issue of 'locked platforms'. Future research should address the extent to which disinformation and fake news (in all its definitions) continue to thrive online, and whether there are monitoring capacities in place so that its overall consumption may be measured, persuasive value judged and wider societal implications discerned.

## **Introduction: Influence campaigning in political spaces online and the question of persuasion**

In reviewing the scholarship surrounding so-called fake news, one would out of necessity make a distinction between the dominant work on the art of 'influence campaigning' and 'computational propaganda' online and the consequences to date for its consumers, but also the few findings, often journalistic, in the relatively understudied case of the Dutch political space online, both on the web as well as in social media. Much work has been undertaken on the influence of Russian (and Russian-style) influence campaigning in the US, and the presence or absence thereof during elections in Germany, France, Sweden and elsewhere. With respect to the Netherlands, the case studies have been reserved to the 'early' Russian influence campaigning around the downing of the MH17 Malaysian airliner (beginning in 2014) and the suicide bombings in the Brussels airport and metro (2016), discovered through the use of Twitter data sets of Russian trolls, or influence campaigners. Other work has been performed on the existence of home-grown troll networks, operating in a predominantly right-wing media sphere, that are at times 'pro-Russian' but do not seem to have had 'foreign' input.

Crucially, in the studies and journalistic treatments to date it is increasingly remarked that there has been a shift in Russian disinformation campaigning from inflaming conflict with the West to stirring it within the West. It is also argued that disinformation could be said to be 'Russifying', i.e., the borrowing of so-called Russian techniques by domestic actors. The campaigning, whether foreign or domestic, does more than create narratives that divide; it also employs computational means to inflate and amplify them through bot work, fake following, astroturfing, the creation of front groups and other artificial publicity tactics.

It is also argued that more attention ought to be paid to the rise of extreme and divisive media on social media platforms, where the point is often made that great emphasis is being placed on foreign disinformation when by comparison it performs poorly in European news spheres. The growth of 'hyperpartisan' news and commentary also may be viewed as an alternative fact or knowledge infrastructure, contributing to discussions of a post-truth condition and the contention that established institutions are under threat.

It is of equal importance to examine the critique on 'persuasion', or the extent to which the influence campaigning strategies, fakery and hyperpartisan sources have discernable impacts on their consumers, especially the voters. They appear to be minimal. Indeed, there is a small, but growing literature critiquing 'transfer' models, also known as 'hypodermic needle' or 'magic bullet' theories which themselves could be considered part and parcel of the fake news hype and fascinations with so-called psyops activities such as in the Cambridge Analytica case.<sup>2</sup> Transfer models do not take into account the 'active filtering' of media users or 'phatic sharing', it is argued, whereby one circulates dubious media more to connect with others or for amusement than to pass along substantive information. Such models also would discount hardened attitude, and studies finding that campaigns generally have 'minimal effects'.

---

<sup>2</sup> The Cambridge Analytica case or scandal refers to the illegitimate use of over 80 million Facebook users' information to develop micro-targeted advertising (Cadwalladr and Graham-Harrison, 2018). It prompted US Congressional and UK Parliamentary investigations, and also led to Facebook's tightening its data access for academics and public scrutiny more generally (Bruns et al., 2018).

As for the measures to be taken, the literature both describes and occasionally questions fact-checking and media literacy efforts because of the assumption that corrected information would assuage knowledge 'deficits', for attitudes often remain the same. Nonetheless, among the policy recommendations most frequently put forward are bolstering media literacy initiatives, together with flagging questionable content, manually and automatically, for further scrutiny. Social media platforms are facing regulation, and are asked to address extreme content and create public archives.

One aspect of the literature review relevant to empirical work concerns the methods employed to demarcate 'political space' online for the subsequent study of the scope and impact of fake news, junk news and computational propaganda -- to use some of the terms for the larger phenomenon under study. Under consideration here are largely mixed (quanti-quali) techniques and digital methods from media studies and data journalism. These provide distinctive political space demarcation strategies for the web as well as social media per platform as well as approaches for cross-platform analysis. They query engines and platforms, measure significant political stories (in terms of engagement) and determine construals of dubiousness through news criticism, categorising significantly engaged-with stories into genres such as disinformation, conspiracy, clickbait, hyperpartisan and (automated) amplification. While often practiced on a story level, the determination of dubiousness also may be made through source criticism, according to the extent to which publishers' output repeatedly accords with junk news definitions, discussed in the next section. It is also worth studying how signal-based or algorithmic determinations of fakeness comport with qualitative methods that are based on source (provenance) criticism.

### **Fake news, junk news and computational propaganda**

Historically, fake news could be thought of as linked to particular novel publishing practices both "when old media were new" but also nowadays through social media platforms (Marvin, 1988; Gitelman, 2006). The term 'canard', meaning unfounded rumour or story, refers to the contents printed in the French broadsheets of the eighteenth century; 'scandal sheets' are the British term for the same era of publishing (Darnton, 2010). In the U.S., in particular, "fake news" as a term recently experienced a revival and travelled internationally, in the numerous senses in which it has been deployed historically: "news satire, news parody, fabrication, manipulation, advertising, and propaganda" (Tandoc et al., 2018). As directed towards contemporary social media platforms, the charge of "fake news" and similar terms often has been uttered as a lament after the introduction of new media technologies, where there are concomitant calls for new journalistic standards, as witnessed with the competing tabloids and their sensationalist, yellow journalism in the late 1890s and into World War I as well as the radio and newswire in the 1920s (Opper, 1894; Lippmann, 1922; McQueen, 2018).

With the rise of corporate public relations, the blurring of the distinction between the 'editorial' and the 'advertisement' sent over the wire or into the airwaves prompted the use of the moniker, "news fakers" (McKernon, 1925; Lazer et al., 2018). Similarly, the contents of the early, "unedited" web, populated by self-publishers, and later the blogosphere, were often described as "too fresh to be true", given the speed of news production and the potential for those looking for a scoop to convert unsubstantiated rumor into news (Hall, 2001; Rogers, 2005). More recently, the notion would be routinely deployed by satirical news sources such as *Saturday Night Live!* in the US (Day and Thompson, 2012); in fact, *The Daily Show*, the progressive comedy news program, described itself

proudly as a “fake news program” (Newman, 2010). Parody, it should be recalled, was behind the origination of the most circulated fake news story during the US presidential campaigns of 2015-2016, “Pope Francis Shocks World, Endorses Donald Trump for President” (Allcott and Gentzkow, 2017). While many definitions concentrate on a story’s falseness, fake news have the “trappings of news” through the use of the news apparatus, including the style, formats and containers employed (Laquintano and Vee, 2017; Grinberg et al., 2019). Indeed, narrower definitions of fake news take as their point of departure how the sources “falsely claim to be news organizations” (Tucker et al., 2018).

Fake news also has been deployed politically as a barb against the free press when publishing inconvenient truths, or speaking 'truth to power' (Cary, 1955; Darnton, 2017). Since the mid-nineteenth century, labelling the news media generally and disagreeable reporting specifically as product of *der Lügenpresse* or the lying press is a discrediting ploy or even communication strategy, still practiced today by right-wing social movements as Pegida in Germany, chanting at street rallies *Lügenpresse, halt die Fresse* (lying press, shut your mouth) (Beiler and Kiesler, 2018). It was the German *Unwort des Jahres* (notorious word of the year) in 2014, in the competition organised by TU Darmstadt. Fake news is also a label, used in highly conservative political circles in the US, for particular news sources, notably CNN, MSNBC, *The New York Times*, and *The Washington Post*; the designation is similar, albeit perhaps more extreme, to past portrayals of the agenda-setting 'elite media' in contrast to conservative upstarts as *Fox News* (Marwick, 2018; Tripodi, 2018; Peck, 2019). In this respect, one could call the current situation just the latest fake news scare, or even moral panic (Brennen, 2017; Morozov, 2017).

When discussing the fake news phenomenon in relation to social media and other online sources, researchers at the computational propaganda project at the Oxford Internet Institute (OII) often offer the umbrella term 'junk news', defined as “extremist, sensationalist, conspiratorial, masked commentary” (Howard et al., 2017). Other catch-alls include “problematic information”, “information disorders” and “false news” (Jack, 2017; Wardle and Derakhshan, 2017). Apart from sensationalist, conspiratorial and masked – features that have been a part of fake news ontologies for centuries – the OII definition emphasises another element, extremist, which cuts to the heart of contemporary concern for the phenomenon when studied not only as a practice of media and public opinion manipulation but also a trigger for societal unrest.

With respect to the growing anxiety over fake news as harbinger of unrest, one may refer to the distinctions made between a variety of 'information disorders', as well as the coinage of new terminology that captures excitable, Internet-related media and speech (Wardle, 2018). First, disinformation and misinformation are both false, but the latter is unintentionally so, whilst the former is fashioned for the purposes of intentional disruption and causing harm. A third term, mal-information (a neologism), seemingly borrowed from malware or malicious software categorisations, has been introduced to describe harassment, toxic content and hate speech online (Wardle and Derakhshan, 2017). These are the tools for the so-called 'weaponization' of social media platforms to foment discord through seeding the news and public opinion with divisive content. Indeed, 'extreme speech' is a term that has been offered as a nuancing of the hate speech discourse as it is applied to online toxicity. It is meant to capture a form of charged language and cultural conflict that stops short of hate, and has emerged with social media, defined as “vitriolic exchange on Internet-enabled media” (Pohjonen and Udupa, 2017). Its rise has prompted social media

companies as Facebook, Twitter and Alphabet (owners of YouTube) to expand their content reviewer pools as well as widen their internal mandates to identify and remove more than violence, pornography and hate (Gillespie, 2018). Google also installed a feedback system for its web search to report inappropriate autosuggestions, after reports of queries for the 'holocaust' autocompleting with 'is a hoax' (Solon and Levin, 2016; Hern, 2017).

As with new media technologies of old, social media platforms currently are said to enable the 'supercharging' or the acceleration of the spread of fake news (Bounegru et al., 2018). Two terms have been used to capture the web and subsequently social media as accelerationist media: clickbait and computational propaganda. Clickbait connotes titillating and sensational content and is formulaic in its presentation, often containing numbered lists (sometimes referred to as a 'listicle') as well as an 'information gap' that sparks curiosity, e.g., 'twenty things you should not do when visiting Japan'. Facebook, in seeking to identify and downgrade clickbait in its news feed, defines it as "a posted link with a headline that encourages people to click to see more, without telling them much information about what they will see" (O'Donovan, 2018). Generally social media companies seek to operationalise substantive definitions into computational signals. Thus, to Facebook, brief attention (or short 'time-on-site') is a signal of clickbait, for readers, having been lured in to the "junk food of content consumption", are subsequently dissatisfied with the low-quality content, and leave the page quickly (DeAmicis, 2014). Clickbait, often innocuous, has been combined with divisive content, however (Burger and Schenk, 2019). 'Extreme clickbait' was a part of the the story behind the allegedly apolitical Macedonian teens based in Veles, who used "spammy techniques" in optimising pro-Trump sites to make money, in the run-up to the US presidential elections of 2016 (Silverman and Alexander, 2016). Follow-up reporting has sought to debunk that narrative, finding that the clickbait campaign was orchestrated by political operatives (Wendling, 2018; Silverman et al., 2018).

Computational propaganda, the second term, refers to the "the assemblage of social media, autonomous agents and algorithms tasked with the manipulation of opinion" (Neudert, 2017). The breadth of the definition is intended to capture the bots that amplify content, the advertising platforms that enable micro-targeting and personalisation of influence messaging, and the click farms that inflate the follower counts and engagement scores, granting posts higher "vanity metrics" and thus greater symbolic power through fake support (Rogers, 2018a). For computational propaganda, bots increase the spread or reach of the posts and inflate their metric counts (Woolley and Howard, 2016). "Low-credibility content" is spread disproportionately by 'social bots,' which refer to bots or autonomous agents tasked with influencing discussion and public opinion; such a finding has led to calls for curtailing their use (Shoa et al., 2018). As a part of the 'assemblage' of actors and software practicing computational propaganda, the work of software-assisted, political operatives has come under scrutiny, especially in the run-up to elections. Sock puppets, assuming the false identity of a grassroots organiser or a concerned individual, create and circulate political content, organise events and mobilise audiences, making interventions in the physical world through hashtags, internet memes and Facebook events (Mina, 2019). 'Front groups' or even faux 'hashtag publics' also mobilise followings and organise demonstrations (see Table one); one notorious case concerned an anti-Islam protest and counter-protest in Houston, Texas, in 2016, where both groups were mobilised by Russian campaigners operating under the names of the Blacktivists and the Heart of Texas, respectively (Shane, 2018).

A related term for fake content insertion for political ends is astroturfing. It is the artificial seeding of newspapers and other content providers with political (or corporate) advertising disguised as genuine citizen concern. Such content is a different category than sponsored political content, where there are regulations that mandate labelling it as 'paid for by' a particular candidate or campaign (Vaidhyanathan, 2017). Nonetheless there have been calls to have 'masked' political content unmasked and marked as sponsored, however much in the recent case of a pro-Brexit group, Britain's Future, investigative journalists were long not able to unearth the funding source, despite the transparency of its being labelled.

Particular forms of native social media advertising have prompted the calls for further public scrutiny of political ads, and also perhaps an expansion of the definition of such. 'Dark posts' (aka 'promoted posts') on Facebook refer to micro-targeted advertisements, without a referral page anchoring the content for further investigation (Bump, 2017). Used by political operatives, including foreign influence campaigners, in the US in 2014-2017 and beyond, such campaigning tactics assemble 'keyword publics' algorithmically by querying the Facebook advertising platform for words such as 'second amendment' or other pro-gun terminology and sending advertisements to the news feeds of the tens or hundreds of thousands of those users determined to have such an interest (Angwin et al., 2017). These publics are targeted not so much because they are persuadable voters but rather to have them circulate and amplify messaging.

<b>2016 Fake rallies planned and promoted</b>		
<i>Date</i>	<i>Fake rally</i>	<i>Location</i>
25 June	March for Trump	New York
9 July	Support Hillary. Save American Muslims	Washington, D.C.
23 July	Down with Hillary	New York
20 Aug.	Florida goes Trump	Several Florida cities
2 Oct.	Miners for Trump	Several Pennsylvania cities
12 Nov.	Show your support for President-Elect Donald Trump	New York
12 Nov.	Trump is NOT my president	New York
19 Nov.	Charlotte against Trump	Charlotte, N.C.

Table one: Overview of 2016 Fake rallies planned and promoted, as listed in the US indictment of 13 Russian nationals concerning foreign election interference. Source: Parlapiano and Lee, 2018.

Apart from particular social media advertising products such as dark posts, other formats have been identified as energizing publics with divisive messages. 'Image macros' are photos with two lines of text, one opening and one closing line, that are a popular format for political messaging on Facebook and have been among the most shared and otherwise most engaged-with content on the platform (Renner, 2017). Indeed, in the data analysis of the most shared posts of the 'fake' (or astroturfing)

activist group pages set up by the Russian Internet Research Agency (Blacktivists, United Muslims of America, Being Patriotic, Heart of Texas, Secured Borders and LGBT United), the image macros and other meme material scored particularly well (Chen, 2015; Albright, 2017; Timberg, 2017).

### **Russian influence campaigning, Russification and the 'hyperpartisan' style**

'Dark globalization' is a term put forward by the historian Timothy Snyder to refer to how knowledge of western societal problems provides opportunities to influence campaigners from abroad, or Russia in particular (2018). In the US Snyder refers to the complex of race, gerrymandering and the electoral college, and the capacity to target voters in specific geographical areas (such as counties in 'swing states') with racist political messaging that amplify or provide 'oxygen' to viewpoints. There have been detailed analyses of the Russian influence campaign of 2014-2017 commissioned by the US Congress, both of which benefited from data provided by Facebook, Twitter and Alphabet (Google) that previously had not been made available for research (Howard et al., 2018; New Knowledge, 2018). They are a part of a litany of literature that has appeared since the commissioning by governments to study the 'tactics' of the influence campaigners as well as the contemporary art of propaganda and the development of counter-narratives more generally. These studies also have led to recommendations concerning how to combat the effects.

The New Knowledge study (and its coverage) emphasise the collective cognitive dissonance that effective propaganda achieves, introducing (and popularising) language from intelligence and counterintelligence work. Among the goals of the propagandists is to create "a wilderness of mirrors", originally a phrase from a T.S. Elliot poem but mobilised by the intelligence community (Holzman, 2008). It refers to an environment where truth (and its establishment) are no longer self-evident (Groll, 2018).

To achieve that goal, New Knowledge argues, one particular tactic is the creation of a similarly termed "media mirage," or "interlinked information ecosystems designed to immerse and surround targeted audiences" (2018: 42). They are enveloped in an information "cacophony," where stories from the press are repurposed, and given another author ('blacktivists'), interpretation and tone. Here is one example, taken from an original newspaper story about how an "11-Year-Old Texas Boy Invents Device to Prevent Hot Car Deaths" (Dahlgren and Arkin, 2017). It was reworked as follows: "[T]hese are stories of Black children the media don't want you to see"; "White people invent tools for killing, this Black child is inventing a tool for saving lives" (New Knowledge, 2018: 62). The divisiveness and the effectiveness ascribed to the sample post derives not only from the transformation of the feel-good news story into a contrived in-group and out-group divide based on race. Note, too, the format used; the second sentence is a two-liner, to be cast into an image macro or meme, the popular format for sharing and further circulation of grievance, outrage as well as mockery. The story also brings together categories of fake news. It is both clickbait as well as extreme content (be it hyperpartisan or junk), as it invites the consumer to read more about the grievance. It is also packaged to be shared.

The purpose of such campaigning is to sow discord and enmity, but it is only one of a variety of tactics where the overall goal is to remove a sense of a collective and shared experience of the world, as analysts have phrased it, and reify group formation (Gessen, 2018). Apart from the creation of a media mirage, the other tactics listed are as follows: "targeting, asset development, cross-platform brand building, memetics, inflecting a common message for different audiences,



narrative repetition and dispersal, repurposing and re-titling pages and brands, manipulating journalism, amplify conspiratorial narratives, sow literal division, and dismiss and redirect” (New Knowledge, 2018: 2). With respect to social media, as discussed above, targeting could refer to the audience segmentation available in platforms for advertising purposes, and memetics to the use of both the image macro to formulate a punchy message as well as to build the meme as an additive content container for narrative reinforcement.

It is worthwhile to mention that the expert studies are snapshots, but these as well as subsequent reporting have pointed to the 'ongoing efforts' of the influence campaigners, and its global spread. While social media companies -- since the Cambridge Analytica and fake news scandals -- have become more active in identifying and suspending accounts of known Russian and other state-sponsored trolls (e.g., Iran), similarly named accounts are active and can be traced to known networks of political operatives (New Knowledge, 2018; FireEye, 2018). New accounts are continually made (Vaidhyanathan, 2018); the Chief Technology Officer at Facebook speaks of “blocking more than one million fake accounts every day, sometimes just when they are created” (O’Brien, 2019). The percentage of influence campaigner accounts in that large number is not given.

Recently, there has been growing concern not only about the ongoing efforts of Russian influence campaigners but also the uptake by other groups (or “domestic actors”) of the so-called 'Russian playbook' (Frenkel et al., 2019). Journalistic coverage was prompted by the announcement by Twitter that prior to the US Congressional elections of 2018 it removed accounts of Americans posing as members of state Republican parties (Harvey and Roth, 2018). Facebook also announced that hyperpartisan pages on both sides of the political spectrum in the US would be removed. Discussions of the 'Russification' of online political campaigning also historicised disinformation, pointing to the classic examples, such as the claim that the HIV virus was the leaked product of a US bioweapons lab; it was planted in news outlets beginning in 1983 by Soviet *dezinformatsiya* campaigners in 'Operation Infektion' and ultimately spread four years later to US TV news (Boghardt, 2009; Ellick and Westbrook, 2018). Comparing the time span of such a news spread to the dynamics of reach in the hybrid media system nowadays is how one may describe how the 'platform press' has supercharged fake news (Chadwick, 2013; Bell and Owen, 2017).

In a well-cited article in the *New York Times*, Facebook, as a leading example of the ‘platform press’, was described as a “totally insane, unintentionally gigantic, hyperpartisan political-media machine” (Herrman, 2016). The author spends some time describing the manner in which Facebook mixes posts in its news feed from both family members and faint acquaintances, but also discusses the presence of upstart media organizations and self-styled advocacy groups that only exist online, many only in social media. Most are described as ‘hyperpartisan’. These sources populating the platform with content are defined as "openly ideological web operations". They also are successful, not just because more extreme and sensational content spreads faster than more sobering truth (Vosoughi et al., 2018). It is also because they employ formats that engage large numbers of users, and learn from their engagement and reach. 'Operating' in a continuous feedback loop of metrics data, posts are optimised to perform well in social media. The performance measures are based on the virality of posts, and those that work well are emulated. There are particular formats as well as styles that drive engagement. Memes and clickbait such as listicles, cliffhanger headlines and human interest stories are among the formats used, as mentioned above. The hyperpartisan style has a variety of substantive features, not all of which are equally applied, but many appear to be working

well. Often anti-establishment as well as positioned as against or in competition with the truth-seeking and fact-finding of the mainstream media, the media operations post stories that are alternatives. These alternatives may be interpretations, facts and editorial commentary on events. They become media layers on the news. The presentation is often edgy, both in terms of being knowledgeable on trend but also sharp in tone. The posts are regular, and as such are part of the permanent updating culture, providing an competing 'feed' about what is happening in the world and in media.

### **The post-truth condition**

There is a series of contemporary utterances that have contributed to public discourse about a post-truth condition. One is the satirical notion of 'truthiness' (Colbert Report, 2005). Developed as political news commentary and comedy, it refers to having the appearance of being true, but without evidentiary basis. Another – 'alternative facts' – is of more recent, political coinage. It refers to the insistence by a member of the Trump administration that the number of attendees at the presidential inauguration in 2016 was higher than reported and measured by crowd science (Still, 2017). The clarification of the meaning behind 'alternative facts' is more to the point: "additional facts, alternative interpretation" (Nuzzi, 2017). Compared to truthiness, here facticity does not derive from appearance or perhaps authority but rather from other fact-making.

In response to what is sometimes considered first-order objectivity battles, or disputes over matters of fact (Margolis, 1995; Latour, 2008), newspaper reporting with such headlines as "Here Are the Real [Facts]" as well as the work by fact-checking bureaus and initiatives are contesting fact claims with increasing urgency (Fandos, 2017). These are public debates about facts, inputs into which include fact-checking, a common practice of journalists and university research groups seeking to confirm the basis behind particular statements by politicians and others (Graves, 2016). Recently, scholarship on the effectiveness of fact-checking has developed in at least two directions: the extent to which fact-checking corrects the record as well as factual beliefs, and whether it changes attitudes (Barrera et al., 2017). Both are part of the decades-long discussion and critique of the 'information deficit' and 'diffusion' models, which challenge ideas that providing correctives clears up controversies (Wynne, 1991; McNeil, 2013).

In the fake news scholarly discourse, it has been found that there are distinct audiences for 'alternative facts' and 'fact-checked facts' (Bounegru et al., 2018). Whilst there may be a correction to the record, the original audience may not have been exposed to it. Fact-checked stories also have similar circulation patterns to alternative facts; they are forwarded to like-minded audiences (Shin and Thorson, 2017). Though it does not tell the entire story about exposure, both the original as well as the fact-checking publications are outlets with distinctive audiences or subscriber bases, with fact-checking newsletters often with smaller, specialty circulations, though their visibility may increase as they are built into the Facebook interface. In the other strand of work, it is asked, does exposure to fact-checked facts change factual beliefs as well as attitudes? Here one set of findings is in keeping with the critiques of the effectiveness of fact-checking and the information deficit model more generally, for respondents saw their factual accuracy 'improve', but their attitudes remain unchanged (Nyhan et al., 2019). Fact-checking, however, could be understood as a documenting process that corrects the record by capturing a dubious story and committing it, and its debunking or exposure, to searchable databases and other media.

The post-truth condition, though, has been described as a competition with respect to not first-order but second-order objectivity. In such a circumstance there is a rise of competing regimes of truth (Fuller, 2018). Expertise becomes 'sectarian' (Turner, 2001). The idea of the media mirage (evoked to describe effective disinformation campaigns) does not in itself create a competing truth regime or infrastructure. Rather, it introduces noise into an infrastructure. But when propagandists, or in a different reading of the contemporary situation, a right-wing media ecology, create an alternative news and information infrastructure, those efforts fit with descriptions of the post-truth condition (Benkler et al., 2017; Sangerlaub et al., 2017).

In other words, post-truth is a term that should not be construed as signifying hoodwinked (or radicalised) consumers, or the 'wholesale cheapening' of fact-making (Sismondo, 2017). Rather, in asking whether 'we can have our facts back', the debate concerns whether (or when) publics can agree on the 'facticity infrastructure' or even the modernist project of knowledge institutions (Marres, 2018). As a case in point, there are ideologically distinctive alternatives to Wikipedia (such as Infogalactic, Metapedia and Conservapedia), producing encyclopaedias challenging not only what is known or settled fact, but also the sources rooting it (Fitts, 2017).

### **Elections, disinformation, and the Dutch case**

Three recurring topics are often discussed in the news and (commissioned) research on disinformation and fake news in the Dutch context. First of all, of particular concern are Russian trolls and their spreading of disinformation in the Netherlands. Secondly, there are the (non-Russian) fake accounts and fake fans that that inflate the popularity of a campaign or a prominent figure, granting them greater symbolic power. And thirdly, publications are addressing its discernibility and possible countermeasures. How to recognise it and combat it? Each of these discussions is often set against the backdrop of a changing news media landscape, whereby mainstream news is increasingly competing with more tendentious and hyperpartisan outlets, and digitization is leading to user-driven and algorithm-driven personalization. That may narrow the horizon of news that users encounter and perhaps increase fringe consumption, though in empirical studies such has not been found (Wieringa et al., 2017). Comparisons of the Dutch situation are also drawn with the US.

While digitization may be changing how people consume news, a study of online news behavior, disinformation, and personalization of the news by Rathenau Institute stresses that in the Netherlands, the traditional news media still hold a firm and stable position in the media landscape (van Keulen et al., 2018). The study also finds that there is not (yet) widespread algorithmic personalization in Dutch media sites. And, in stark contrast to the current situation in the US, Dutch news consumers tend to use a variety of sources and have trust in the traditional news media (and less so in social media). Lastly, the report underlines that the Netherlands does not have such a particularly polarised media landscape as the US.

Overall, there is a strikingly moderate tone of voice in the literature, both in news reporting and research reports. Since 2016, several studies have looked at disinformation practices in the Dutch political landscape, and each of them has concluded that neither is there any large-scale disinformation activity in the Dutch media nor does disinformation have a significant impact on Dutch citizens. However, in the Summer of 2017, Wilfred Rietdijk, a Dutch general and national security advisor, announced in an interview with Dutch newspaper *de Volkskrant* that the Netherlands could no longer deal with the digital threat (van Zijl and Modderkolk, 2017). A "landslide

of fake news”, as the subsequent tabloid headline read, would lead the country into chaos and division (Jonker, 2017). Including breaches and intrusions in his threat assessment (thereby widening the scope beyond disinformation), Rietdijk explained how Dutch companies are “in the line of fire” from “thousands of hackers from Russia, China, and countries such as Iran and even Sudan” (van Zijl and Modderkolk, 2017). The general is not the first to warn of Russian interference in the Dutch online space, though case studies were lacking, at least in the public domain.

#### *Russian trolling and its perceived insignificance in the Netherlands*

When the Minister of Internal Affairs, Kajsa Ollongren, warned the Dutch government of Russian disinformation in the Netherlands, she initially was criticised for not having compelling examples (Pleijter, 2017; Kist and Wassens, 2018a). Two journalistic studies that have looked into Russian tweets have found activity in the Dutch online realm, however. A study by *NRC Handelsblad* mined 200,000 tweets from Russian Internet Research Agency (IRA) accounts and found disinformation campaigning beginning in 2014 and another spate in 2016. The weekly magazine *De Groene Amsterdammer* combined the *NRC Handelsblad* data with larger collections of Russian troll accounts, made available on the American public opinion analysis website, *FiveThirtyEight* as well as the lists published by American Congress (van der Noordaa and van de Ven, 2018a). Both studies found a peak in trolling activity after the downing of MH17 in July of 2014. The *NRC Handelsblad* study finds that Russian trolls posted 57,500 tweets, most of which were in Russian and aimed to influence public opinion in Russia and Ukraine, and only four of the tweets were in Dutch (Kist and Wassens, 2018b). The study by *De Groene Amsterdammer* confirms that most tweets on MH17 were in Russian, but finds more mentions of Dutch “conspiracy theorists and activists”, indicating a shift from challenging Western narratives (for Russian-speaking audiences) to seeking to stir conflict within the West.

A second event revealed more coordinated Russian troll activity in the Dutch language Twitter space (in Belgium and the Netherlands), and a further example of striving to foment unrest, albeit unsuccessfully (according to engagement measures) (van der Noordaa and van de Ven, 2018b). It concerned the spreading of anti-Islam content directly following the terrorist attacks in the Brussels airport and metro in March 2016, and in the two years after the attacks. This anti-Islam ‘campaign’ involved about 950 tweets in the Dutch language that were circulated by some 150 IRA-related accounts. These tweets were rarely retweeted, however. In the event, Russian trolls are more successful in the Netherlands with the circulation of English-language content. While these tweets are not related to Dutch issues and focus on for instance the US elections, they have been shared widely by over 6,000 Dutch Twitter users with a total of 9.5 million followers (Kist and Wassens, 2018a).

Perhaps counterintuitively, there was only minimal Russian interference with the Ukraine referendum in the Netherlands in April of 2016 (NOS, 2017). There was the Russian video capturing fake Ukrainian far-right militia members threatening terrorist attacks in the Netherlands and burning a Dutch flag, but it was readily recognised as propaganda (Bellingcat, 2016). Otherwise, only a handful of tweets propagating a 'No' vote were found in the larger set of tweets under study (van der Noordaa and van de Ven, 2018a).

The *NRC Handelsblad* concludes its work on the Twitter data set by noting that it is possible there is larger scale Russian activity in the Netherlands; it should be studied beyond just Twitter to include

other platforms with known troll activity, such as Facebook, Instagram, YouTube and Reddit. Indeed, especially after Trump's victory in the US presidential elections of 2016, many news outlets pointed towards Facebook. As discussed in some detail below, a study by *BuzzFeed News* compiled the most engaged-with posts in the nine months prior to the elections and found that fake news during that time was circulating more than mainstream news. Journalists from the *NRC Handelsblad* replicated the study's general method for the Netherlands, but with a more narrow definition of fake news. They determined that the one hundred most-shared political news articles from January and February of 2017, in the run-up to the Dutch general elections, did not contain fake news (Kist and Zantingh, 2017). Certain articles could be considered misleading or biased, they thought, for they exaggerated news facts or took them out of context. The themes that were most resonant during the campaign period in the Netherlands were immigration, Islam and Geert Wilders.

#### *Dutch fake followers and trolls*

Until November of 2017 much of the reporting has insisted that the Netherlands -- and the Dutch elections in particular -- have been largely unaffected by disinformation or fake news. Much of the news coverage that speaks of it concerns 'fake followers'. For instance, in 2015, there was a small scandal about Geert Wilders concerning a dubious increase in his followers on Twitter. Indeed, when Twitter addressed the issue of fake followers and follower count inflation through a mass removal of suspect accounts in 2018, Wilders as well as other Dutch politicians (including from the political party Denk) saw their metrics decline (NOS, 2018). In perhaps the most well-known case, the Dutch singer-songwriter Dotan was found to have a fake following of 140 user accounts, which were used between 2011 and 2017 to like the musician on social media, edit the Wikipedia article on the artist, request his songs at radio stations and circulate heartwarming stories about him across social media platforms. One of the profiles declared how Dotan's music helped her through a period of grief after a miscarriage; another tells how Dotan welcomed one fan's terminally ill brother in a meet-and-greet, throughout which the singer held the boy's hand. Both testimonials were false, as reporters of *De Volkskrant* found and Dotan later confirmed (Misérus and van der Noordaa, 2018a; 2018b).

In 2018 the first large-scale global study of computational propaganda was published, examining organised social media manipulation such as the use of fake followers in 48 countries, including the Netherlands (Bradshaw and Howard, 2018). The study describes the different computational tactics employed not so much by Russian influence campaigners but by political parties to influence voters and the elections.<sup>3</sup> It was found that the use of social media as an infrastructure for the spread of propaganda and disinformation has become widespread. Under examination is "cyber troop activity," defined as "government or political party use of social media to manipulate public opinion" (Bradshaw and Howard, 2018: 9).

---

<sup>3</sup> The research conducted a content analysis of news articles reporting on cyber troop activity in a sample of 48 countries, supplemented by an in-depth secondary literature review. To collect the news articles, the researchers used the following keywords in combination, in queries across Google, Yahoo!, Bing and LexisNexis: astroturf\*; bot; Cambridge Analytica; Facebook; fake; fake account; disinformation; government; information warfare; intelligent agent; military; misinformation; persona management; pro-government; propaganda; psychological operations; psyops; social media; sock puppet\*; troll\*; Twitter (2018:8).

While in more authoritarian regimes, social media manipulation fits into larger scheme of voter suppression and election rigging, in “emerging and Western democracies, sophisticated data analytics, and political bots are being used to poison the information environment, promote skepticism and distrust, polarise voting constituencies, and undermine the integrity of democratic processes” (Bradshaw and Howard, 2018: 5). The tactics described include the use of three kinds of fake accounts. First, there is the creation of online commentator accounts that attack and troll genuine users, spread divisive content, or “[divert] conversations or criticism away from important issues” (Bradshaw and Howard, 2018: 11). A second tactic entails automated accounts or political bots to automatically flood particular hashtags, and astroturf by faking a follower base. The bots also troll genuine users by reporting them and flag organic content thereby having both suspended until a human moderator checks them. A third tactic is the use of hybrid accounts, which are those that make use of automation (for the sake of speed and convenience) but are actively curated by human users, who commonly manage multiple fake accounts or sock puppets. This type of fake account is difficult to recognise, and thus to combat. The study finds that automation is the most recurring tactic, seen in 38 of the 48 countries that were shown to operate fake accounts.

Besides the use of fake accounts, other strategies involve the use of political ads and the involvement of search engine optimization and activity on chat applications and across social media platforms. Where Twitter is proven to be the platform most friendly for automation, the study finds “cyber troop activity on chat applications or other platforms (Instagram, LINE, SnapChat, Telegram, Tinder, WeChat, WhatsApp)” in one-quarter of the countries under study (Bradshaw and Howard, 2018: 13). In the European countries in their sample, they find a distinct fake news footprint per country. In Germany, junk news is rather marginal and was mostly circulated by right-wing political actors during the 2017 federal elections. In Italy on the other hand, a large and active “ecosystem” of it is connected to populist political forces such as the Lega Nord (Northern League) and the Movimento Cinque Stelle (M5S, 5 Stars Movement), which were at work during the 2017 constitutional referendum and the elections of 2018. Here, fake news connects national politics to Euroscepticism, conspiracy theory, aliens and pro-Putin propaganda. In the Netherlands, this analysis also finds that disinformation revolves around politician Geert Wilders and in particular the spread of his anti-Islam video, which was broadcast on television and shared in social media in the lead-up to the 2017 Dutch national elections. In particular, the study finds that automated accounts have amplified Geert Wilders' campaign hashtags.

These results match the findings in a study that looked at troll-like behavior on Twitter, leading up to the 2017 Dutch general elections, where sock puppets were found (Bounegru et al., 2018). The study collected over 500,000 tweets mentioning at least one of the Twitter accounts of the 28 political leaders a month before the 2017 Dutch general elections. To retain the users that demonstrated troll-like behavior, it narrowed down the set to retain only the 25 users who consistently targeted one or more political representatives.<sup>4</sup> The analysis showed that there was a notable asymmetry in the distribution of targets of troll-like behavior and sockpuppetry across the political spectrum, where left-wing politicians are most often targeted by negative mentions, while right-wing politicians receive support. Troll content extended to reputable news sources which cited it at least thirty times. Among the cited troll accounts were fake news organisations with names as ‘Today in Syria’ and ‘WorldNewsPolitics’, political parties (including multiple fake accounts for the Republican

---

<sup>4</sup> By @mentioning them at least 100 times in a one-month period.

party in Tennessee) and concerned citizens, most of whom were fiercely pro-Trump and anti-Islam (Kist and Wassens, 2017). In another analysis by the *NRC Handelsblad*, a Dutch political party (DENK) also exhibited troll-like behaviour, including sockpuppetry on both Twitter as well as Facebook (Kouwenhoven and Logtenberg, 2017).

While Dutch news consumers have been found to use a variety of news sources, the Netherlands also has a steady junk news diet (Burger et al., 2019; van der Poel, 2019). From 2013-2017 Dutch Facebook users consumed more low-quality, commercially driven clickbait than mainstream news, as was found through engagement scores. As may be expected, there is also relatively more clickbait on Facebook than quality news.

The consumption and forwarding of clickbait, extreme clickbait as well as other problematic information extends also to politicians and public figures. One Dutch researcher, Peter Burger, has a collection of instances when Dutch politicians have retweeted anti-semitic or otherwise disturbing content. In one example, a video purporting to show 'Muslims vandalising Christmas market in Lithuania' was actually a recording of an event that took place in Baltimore in the US (Burger, 2016).



Figure 1: Cartoon that ridicules the fake news taskforce, stating: "internet trolls are best countered by internet hobbits." Source: Reid et al. (2018).

#### *Recognizing and countering disinformation in the Dutch online space*

Various initiatives aim to detect and counter disinformation in the Netherlands and on an EU-level. The EU taskforce (*East Stratcom Task Force*) against disinformation was heavily criticised in the Netherlands after its project *EUvsDisInfo* mistakenly categorised articles by *The Post Online*, *GeenStijl* and *De Gelderlander* as disinformation (van Keulen et al., 2018; Heck, 2018). (Figure 1 shows a cartoon about the fake news taskforce, stating internet trolls are best countered with internet hobbits.) In a sense the dispute stemmed from misreadings of single stories, perhaps without an appreciation of how settled some of the sources are in the Dutch media landscape, despite their tendentious style (in the case of *The Post Online* and *GeenStijl*). For its part, *De Gelderlander* had taken over nearly verbatim a Russian storyline concerning the perpetrator behind the downing of the MH17, but did attribute it to its original source in a barebones reporting style. The flagged cases were removed from the *EUvsDisInfo* site after complaints by the Dutch media organization *Persgroep* (*EUvsDisInfo*, 2018).

Fact-checking as a journalistic practice has taken hold in the Netherlands. Many newspapers have implemented (or revived) a fact-checking section, often dedicated to checking statements made by

political figures in interviews in newspapers or TV shows. There are also websites such as Hoaxmelding.nl and Nieuwscheckers.nl that compile lists of instances of false news on Facebook and elsewhere. For their study of disinformation, Rathenau researchers analyzed these lists, comprising respectively 140 on Hoaxmelding (collected between 1 February 2014 and 18 December 2017) and 166 on Nieuwscheckers (between 3 February 2017 and 5 January 2018) (van Keulen et al., 2018). They found that the items on the list of Hoaxmelding involved examples of unfounded warnings (65), polarizing disinformation (32) and fake crime news (31). Additionally, there were several examples of clickbait, benign as well as malicious. The content steers users to advertising, “like-farming” and phishing sites (van Keulen et al., 2018: 38). Such posts contain human interest stories that are “painful on a personal level” (van Keulen et al., 2018: 45). The researchers found that only 25% of the disinformation concerned political content and most clickbait serves a commercial goal, rather than a political one. On the list of items collected by Nieuwscheckers, the Leiden University-based initiative, less than half was found to have political content. Within the full set, the researchers found six examples of polarizing content. Otherwise, many of the posts concern factually incorrect, public statements by politicians, the investigation of which is how fact-checking is conventionally practiced.



Figure 2: “Detected and eliminated” fake news, with a warning issued by NU.nl and Nieuwscheckers. Source: NOS (2017a).

Fact-checking now extends well beyond unpacking politicians’ statements, and Facebook has entered into partnerships with many bureaus around the world, including the Netherlands, to explore and catalogue dubious content. In 2017 Nieuwscheckers partnered with Facebook and NU.nl and celebrated their first collaborative “successful detection and elimination of fake news” that year when they flagged a tabloid-style, human-interest post about an Australian newborn weighing 20 kilograms (see figure 2). In February of 2019, however, Nieuwscheckers withdrew from the Facebook fact-checking initiative because of liability risks (Kist, 2019). Nu.nl continues to work with Facebook on fact-checking, and they do so on a paid basis, an issue raised repeatedly in the context of journalists’ being asked to address an issue of Facebook’s making on a voluntary basis.

The effectiveness of fact-checking as a strategy in the Netherlands is a different question. As mentioned above, fact-checks and fake news often have separate publics, and fact-checks may lead people to fake news, rather than away from it. A recent study in the Netherlands found that even



when many people would agree with a fact-check, they are not interested in reading the fact-checking article, prompting the scholars to advise journalists to make the fact checks an engaging read (Hameleers and van der Meer, 2019). Another strategy to counter disinformation concerns a strand of media literacy that involves developing skills to recognise fake user accounts and disinformation. One is on a source level, the other on a story level. The *Field Guide to Fake News* provides a method for the detection of trolling accounts by looking at their friends, or their profile information (Bounegru et al., 2018). There are also courses and training modules for fake news detection and fact-checking, such as those given by Radio Netherlands (RNTC, 2019). The other format is the fake news quiz, such as those by *de Volkskrant* (2016) and the *Guardian* (2016), as well as the *New York Times* 'deceptive Facebook post' test (2018). These quizzes make it clear how challenging it is to recognise fake news. The Dutch serious game, titled *Slecht Nieuws* ('Bad News'), invites players to create fake news and by doing so gain insight into the strategies behind it and become more astute in its recognition (NRC, 2018; DROG, 2018). It is part of efforts that study fake news as risk and ultimately seek to inoculate populations against it (Roozenbeek and van der Linden, 2018).

#### *Voting aid applications*

Voting aid applications (VAAs), often called *stemwijzers* in Dutch, are generally websites that describe their purpose as helping undecided voters find the political party that best matches their preferences and positions. As such, in the context of the study of fake news, they could be regarded as a competing persuasion instrument, a preemptive measure against influence campaigning, or even a potential site that may include it, either through parody, hoax or hack. The literature on VAAs takes up the Dutch and Belgian cases, countries that together with Germany, Austria and Switzerland have upwards of half the voter population accessing them prior to elections. The work can be positioned broadly as pertaining to “the impact of internet-based applications on politics” and can be roughly divided into user studies, impacts of VAAs on the voters as well as the methods behind them (Hirzalla and van Zoonen, 2015: 88). To date these online voting aids have not been raised as recommended technology to combat fake news and influence campaigning per se, though they do furnish a rather personalised information experience that may be studied for its ‘influence’ effects, as discussed briefly below.

In studies of their usage, researchers have asked whether VAAs “mobilize the mobilized” (Hirzalla and van Zoonen, 2015). And indeed, while VAAs have a heterogenous user base across demographics, interests, attitudes and behaviour (Vassil, 2011), there is an overrepresented subgroup of younger, mainly left-of-center, urban and well-educated male users who are politically active or knowledgeable. This imbalance could lead to the conclusion that those who would benefit the most from political advice are the least likely to seek it (Ruusuvirta, 2010).

A second set of literature concerns the impact of VAAs and assesses whether they have influenced the voting behaviour of its users, though it is not clear whether the quality, reach and graphical interfaces of the aids affected the extent of the influence. From those surveyed anywhere from 1% to 15% using *DoeDeStemTest* (in Belgium) as well as *StemWijzer* and *Kieskompas* (in the Netherlands) reported having been influenced by the aids (Walgrave et al., 2009; Hirzalla and van Zoonen, 2015). While research has found that the politically knowledgeable and engaged users that are common to use VAAs perceive them as useful, they are also among the less likely users to be influenced by them (Alvarez et al., 2014; Dumont and Kies, 2012).

A third set of literature concerns the methods used by the VAAs. Here there is a distinction between the choice of the policy positions to include in the interactive system and the models underlying the advice. The very selection of the policy positions is a crucial factor in the voting advice given, where another set would lead to other advice (Walgrave et al, 2009). In general, VAAs are found to select policy positions according to their saliency (for the election period), and variability (in that different parties hold different positions) (Hirzalla and van Zoonen, 2015). The editorial process differs, where certain VAAs select their statements solely with experts such as political scientists or journalists (e.g., the Austrian VAA *wahlkabine.at*), while others co-create the formulation of VAA positions, workshopping them with party representatives in the case of the Dutch *StemWijzer*, or with an editorial board that consists of professional experts as well as first and second-time voters in the German “Wahl-O-Mat” (Garzia and Marshall, 2017).

As the voters register their political views, and in certain cases add weight to them, the software calculates the extent to which the voters’ preferences match the respective parties’ and presents its results as a ranked list, bar chart, grid or radar chart. Several studies concentrate on the workings and visual outputs of the different voting aids. Louwerse and Rosema (2014) dissect them by examining how many dimensions are taken into account when ranking the political parties. In their study, a one-dimensional model refers to the ranking the political parties based on the level of agreement with the voter, and presents its findings as a ranked list or bar chart. A two-dimensional model places the political parties’ statements and the voters’ responses on a continuum from left-wing to right-wing, and proposes its match accordingly. The more elaborate multi-dimensional model, employed by the Swiss *smartvote* application, plots the statements and responses onto eight policy dimensions and presents its results in a spider plot that is more complex to read (Louwerse and Rosema, 2014). In a comparative test of these models, researchers took a dataset from the Dutch *Stemwijzer* and found that the different spatial models would lead to very different matches (Louwerse and Rosema, 2014).

As mentioned above, the voting aids are rather popular in a series of European countries, and could be considered not only as another information input but also as one that competes with campaigning. Though the influence (similar to campaigning) may again be minimal, it could be considered as another approach or counter-measure in the discussion of how to address the fake news problem.

### **Fake news may be pervasive, but is it persuasive?**

If one were to divide the current period of fake news studies into waves, it could be argued that the first related to the definitional issues and the production side (as mainly discussed above), whilst the second is increasingly concerned with the study of its consumption (Boczkowski, 2016). In other words, fake news may be seen as “pervasive, but is it persuasive?” (Shaw, 1979). Why do people consume fake news, and do these readers have particular demographics or profiles? Which people deem these stories credible or at least have pass-along value? Are they persuaded or even persuadable? In the US and in a growing list of other countries social media platforms are increasingly a main source of news, and the manner in which they deliver news is different from a newspaper or similar package or container (Gottfried and Shearer, 2016; Poynter, 2019). One receives single stories, rather than an entire newspaper, each shared by someone with whom the social media user has made a connection, most often directly. These can be friends (Facebook),

followers (Twitter), connections (LinkedIn), etc. Stories arrive in the feeds algorithmically, meaning there is a filtering mechanism where certain of them are boosted, based on signals such as activity and increasingly trustworthiness, or the amount of given and measured meaningful engagement between individuals. Put differently, those who are 'close' to the user (by some special measure) are the ones whose stories more likely will be seen (Eslami et al., 2015). Such observations have led to discussions of the re-application of the notion of the filter bubble, a term originally associated with a user receiving personalised (rather than universal) search engine results (Pariser, 2011; van Keulen et al., 2018; Puschmann, 2018). Personalisation, however, has evolved from being the result of the information interactions of one user searching to engagement with an entire social network. As such it shifts the bubble from enveloping the individual to the group; it has prompted 'bubble studies' of not just social media news environments, but those of health, science, fashion and other areas of collective information production, sharing and recommendation (Pedersen and Hendricks, 2014; Hendricks and Vestergaard, 2019). Indeed, fake news circulation and consumption are increasingly experienced as an issue for the environment (e.g., climate change and its sceptics), health (e.g., the anti-vaccination discourse) and a variety of other areas (Kitta, 2018).

Such findings have led researchers to define on the one hand the groups most likely to consume and share fake news together with the dynamics of their bubbles, and on the other the meaning, or sincerity attached to the sharing. In terms of the consumption of fake news, it could be said at the outset that there have been two widely cited findings about their significance from the journalistic arena. One found the most shared stories during the US presidential elections were fake news (see figure 3), and the other that Russian disinformation campaigns had a far greater spread than previously imagined as well as reported in testimony by Facebook before the US Congress (Silverman, 2016; Timberg, 2017). These findings have since been put into a broader context and compared to 'normal' political campaigning and the development of messaging strategies, filtered through news. First, in the event, only a small fraction of the population consumed such 'news' (Allcott and Gentzkow, 2017). Given the limited exposure, the impact, if at all, would have paled in comparison to political TV commercials (Persily, 2017). There is the larger question, however, of whether the messaging would have anything but "minimal effects" (Lazarsfeld et al., 1948). As has been repeatedly found, the net effect of campaigning, albeit by political elites, that persuades the prospective voter is exceedingly low or even zero (Kalla and Broockman, 2017). The aim then is less to persuade than to "rous[e] the enthusiasm of existing supporters" (Panagopoulos 2016).

Though they may have began as symmetrical studies of the right and the left, of the most circulated findings to date about the spread of fake news -- both with respect to the contents as well as its consumers -- ultimately all have overwhelmingly concentrated on the right, be it conservatives and the alt right in the US or other right-leaning, populist right or new right publics in Europe (Bounegru et al., 2018; Benkler et al., 2018). It is of interest to note for starters that both during the US presidential campaigning and thereafter the information spaces or spheres of the right contained far more fake, junk, disinformation or otherwise dubious stories and sources than the left (Faris et al., 2017). Thus conclusions drawn about right-leaning publics sharing information should take into account that they are disproportionately exposed to such information; all else being equal, the right would share more of it (Marwick, 2018). In the empirical studies it was found that the right (most notably Trump supporters) consumed the most fake news. However, there seems to be an older, hard core of its consumers in the US during the run-up to the US presidential elections in 2016 -- "the 10% of Americans with the most conservative information diets" (Guess et al., 2018). These are

heavy media users, and 'available audiences', who have made time to consume media (Nelson and Taneja, 2018). Unlike the majority of the media-consuming public, they are far more likely to read niche rather than only establishment sources. There is, in other words, a normalcy to the consumption by those audiences of fringe materials.

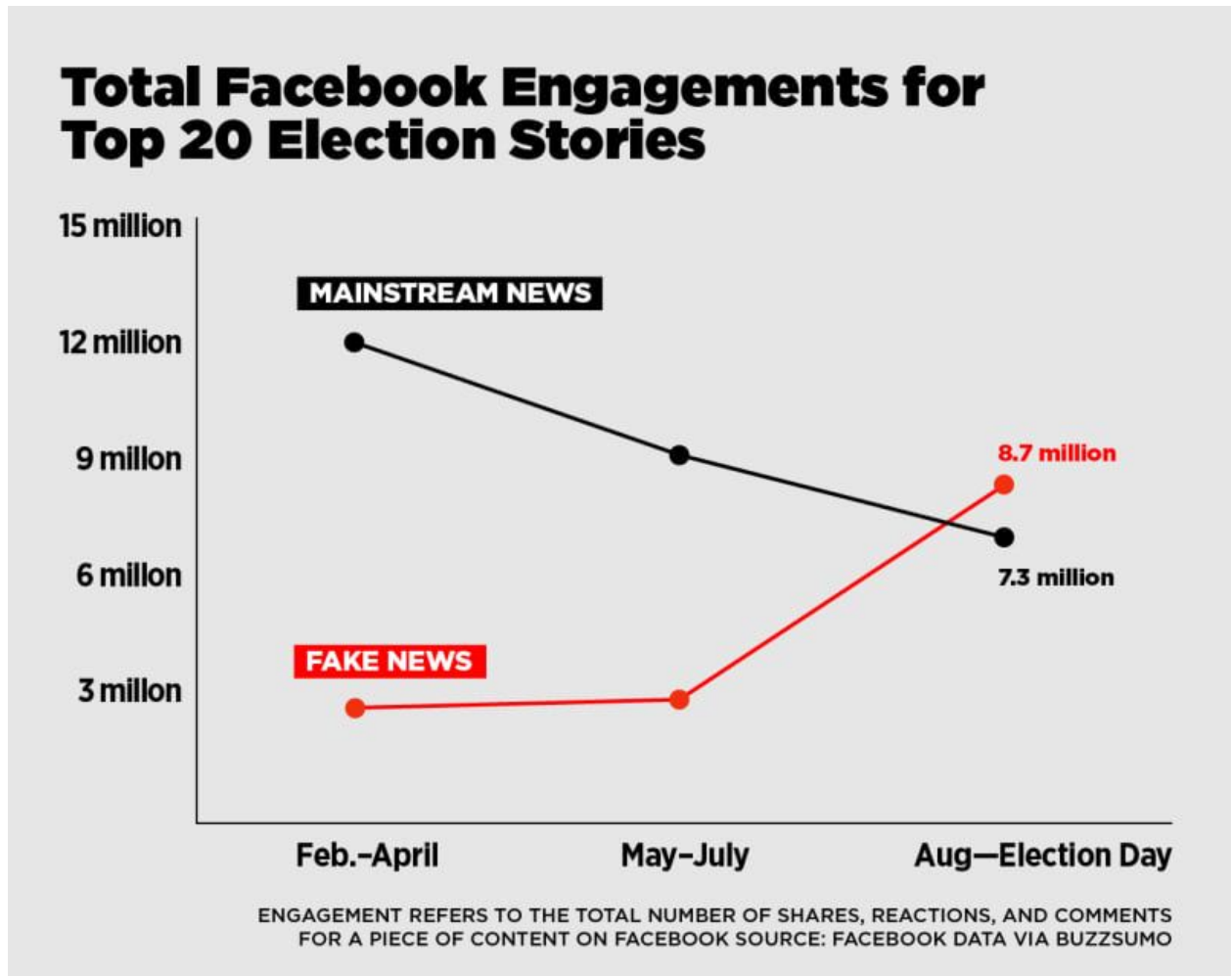


Figure 3. The birth of the fake news crisis, or 'fake news' outperforms 'mainstream news' on Facebook, in the run-up to the U.S. elections in 2016. Source: Silverman, 2016.

The strand of work that considers why users share fake news should be prefaced by the distinction between “earnest and ambivalent” internet users (Hedrick et al., 2018). Much of the scholarship about internet culture has not considered that considerable cultural production and sharing are undertaken not to be part of participatory culture, connective action and other earnest forms of civic culture online but rather for unsympathetic amusement (aka 'lulz') (Phillips, 2015). 'Sharing', a term that has mutated in digital culture from acting in a gift economy to a dominant form of so-called platform capitalism, could have been prompted these days as much by insincerity as by mindfulness (Barbrook, 1998; Belk, 2007; Srnicek, 2017). That is, the rationale for making and sharing could go “either way ... complicating an easy assessment of authorial intent” (Phillips and Milner, 2018: 11). Such a tricky attribution of intent is especially troublesome in the spaces where vitriolic exchange as well as extreme speech and content are prevalent. It is difficult to disentangle whether one is sharing for amusement and to trigger a reaction, or for substantive reasons.

As has been found in the US context, the fake news stories most shared on social media resonate with particular grievances (about the bias of establishment sources) and resentments (concerning economic opportunity) that underlie certain societal divides (Marwick, 2018). Moreover, the stories do not stand alone in a mirror world of conspiracy theory but rather are contiguous with more mainstream conservative news, anchored by Fox News; they are more extreme as well as transgressive in their wording and presentation. Hence the notion of ‘hyperpartisan’, but also there is also reference made to tendentious, anti-establishment sources. Here the Overton Window is appropriately referenced, meaning the bounds of current, acceptable public discourse, and the extent to which extreme speech in hyperpartisan and tendentious sources is moving established norms (Daniels, 2018).

### **Fake news studies: Digital methods and data journalism**

As we come to shortly, one research strategy for measuring the prevalence of fake news story types and sources around national elections is to gauge their presence generally in scoping exercises, but also more specifically in the most engaged-with content in social media concerning elections, political parties, candidates and social issues. A more subtle analysis would examine the top stories for the penetration of fake news narratives, measuring mainstreaming. Moreover, through comparison of engagement with fake news, one also could determine which platforms are most susceptible (or amenable) to hosting and circulating such content. Facebook in particular has been held up as a fake news machine (Herrman, 2016). Empirically, it has been found to host (proportionately) more of it, whether narrowly or liberally defined, than other platforms (Guess et al., 2018).

In order to contextualise such measures it is worthwhile to consider the ways in which the scale, reach and impact of fake news have been studied to date with media analysis, or more specifically digital methods and data journalism. The methods generally could be considered mixed quantitative/qualitative approaches. They often begin in the journalistic arena, with the qualitative determination of the dubiousness of a set of sources and/or stories, and proceed with digital methods that design queries and collect data from platform APIs, media monitoring company dashboards, and social media companies that have furnished lists of banned trolls or user accounts. Indeed, with respect to the dubious source lists, BuzzFeed News' original list of about 20 sources determined to be fake news inform a series of empirical studies (Silverman, 2016; Allcott and Gentzkow, 2017; Bounegru et al., 2018; Marwick, 2018; Grinberg et al., 2019). For studies of the Italian news space, the lists relied upon are from BUTAC, Bufale and Bufalopedia (Fletcher et al., 2018; Butac, 2018). Hoaxwijzer's list of 92 Dutch-language “false news” sites also informs certain of the empirical studies to date in the Netherlands (van Keulen et al., 2018; Wieringa, 2017).<sup>5</sup> But other work, such as the *NRC Handelsblad*'s analysis of the extent of the problem of fake news in the Netherlands in the run-up to the 2017 national elections, looks at the sources afresh, making on-the-spot determinations of fakeness (Kist and Zantingh, 2017). These may conflict with previous listings. For example, Hoaxwijzer lists *De Dagerijkse Standaard* as a “false news” site whereas the *NRC Handelsblad* did not determine it to be “fake news”, but the “extreme right news site” fell among those they called “misleading” because it reported that “1,000 crazy Muslims” had “torched” a church in Dortmund on New Year's eve when instead a firework had landed on its roof causing

---

<sup>5</sup> As other studies also found, the list is dated; as of April 2019, 40 of the 92 sites are offline. It remains useful as a list for older media corpora.

slight damage. The *NRC Handelsblad* determined that it did not meet its definition of fake news as a “fully fabricated story packaged as news”.

Indeed, the question of detecting fake news, on a source or story level, is often placed at the feet of journalists, media organisations and fact-checking bureaus, where credibility and transparency may be rated (NewsGuard, 2019). Masked sources are penalised, for example. As mentioned above, for online stories, the determination of dubious content may benefit, too, from a genre analysis (Lüders et al., 2010). Disinformation, conspiracy, clickbait and (automated) amplification have styles (Rony et al., 2017). Disinformation tends to be a hard counterfactual presentation, conspiracy has multiple characters and plot entanglements, clickbait is a cliffhanger that is often painful on a personal level and (automated) amplification posts at particular intervals and in coordination, as malicious social bot detection projects have found (Ratkiewicz et al., 2011; Bessi and Ferrara, 2016; RoBhat labs, 2017). Other technical signatures of dubious news sites are of interest. For instance, empirical work on the types of cookies and third-party elements in mainstream and nominally fake news sites found distinctive types in each, with the mainstream sites using customised trackers and the other off-the-shelf (Bounegru et al., 2018).

With the lists of fake news sources either in place or determinations still to be made, the next step is to build a media corpus. Following *Buzzfeed News*' method, many undertakings query media monitoring services (such as Buzzsumo, Com Score and Facebook's Crowdtangle) for keywords and/or URLs, in order to build source sets of most engaged-with media and pull in engagement scores per story. Certain of the techniques also include further interpretative coding of stories, including grievance narratives (Marwick, 2018).

Whilst much attention has been directed towards Facebook, and the study of the election-related stories most engaged with on that platform, Twitter is often used as the preferred data source, given dedicated data sets (made available by Twitter or academic researchers) of accounts run by the Internet Research Agency (Farkas and Bastos, 2018). There is a series of studies that rely on Twitter's curated sets as well as on the data robustly collected and shared among data researchers, such as by Clemson University and FiveThirtyEight, mentioned above. In a form of crowd science, the publication on GitHub of the Clemson data set led to numerous studies; in the US widespread disinformation campaigning was found, as is known, but also more niche-targeting of politicians in such states as Maine (Roeder, 2018). As in the Netherlands, discussed above, the data were put to use in other countries that according to journalistic accounts had been previously understudied. For example, in Italy IRA trolls posted numerous pro-populist party tweets in Italian, joining the 'cacophony' or media ecology around the right, as also discussed in the OII work (Fubini, 2018; Fletcher et al., 2018). Twitter is also preferred given the general ease of use of data collection through its streaming and search APIs, intermediate services such as Hexagon Crimson for samples as well as the availability of historical data sets, albeit at a price.

Other approaches (considering consumption and persuasion rather than definition and production) should be touched on that rely on surveys, user data collection and experiments. Numerous experiments have been performed on misinformation (Jankowski, 2018). Recently, for example, a representative sample of the population consents to having their online media consumption passively monitored, and subsequently surveyed thereafter (Guess et al., 2018). Or, there are experiments that show fake news to consumers, and subsequently provide fact-checks to them in

order to determine whether the fact-checks should be 'attitudinally congruent' for them to be persuasive (Hameleers and van der Meer, 2019). In another experiment in the Netherlands, commissioned by the newspaper, *de Volkskrant*, respondents were presented with fake news around one of four topics: vaccinations and autism, MH17, rape incidents in connection to migration, or Sylvana Simons (a politician and media personality) and discrimination. The study tests whether they became less certain about the facts after being misinformed (I&O Research, 2017; Kranenberg, 2017).

With respect to platforms other than Facebook and Twitter, YouTube and especially Instagram appear to be relatively understudied but significant, and Reddit and 4Chan are being recognised as breeding grounds for some of the more outlandish and consequential content such as Pizzagate (New Knowledge, 2018; Tuters et al., 2018). There are platform-specific approaches for building and analysing datasets for Instagram (through queries for hashtags and place names), as well as YouTube, Reddit and 4Chan (Rogers, 2018b; Rieder et al., 2018). Google web search also has invited scrutiny, given the extreme content returned for queries such as the holocaust.

Buzzfeed's work on detecting and analysing fake news on Facebook has been particularly influential in data journalism research and subsequent studies that build upon it, and thus is worthy of mention in some detail (Silverman, 2016). First, the researchers built a keyword list concerning elections (and especially controversial election topics), and subsequently queried those keywords in media monitoring software (Buzzsumo) that returns stories ranked by engagement scores. With the aid of the results, they built a fake news and hyperpartisan website list, which they merged with lists of the same that they curated previously through separate reporting, including on the infamous collection of about 100 websites created by the Macedonian clickbait makers, members of the same family of sites (with the same Google Analytics ID) of WTOE 5 News that created the story about the Pope endorsing Trump, and a collection of hyperpartisan sites (Silverman and Singer-Vine, 2016; Silverman et al., 2016). They also curated a list of some 20 mainstream news sites.<sup>6</sup> (All the accompanying data BuzzFeed also made available through online Google spreadsheets, in keeping with emerging standards in data journalism.) The engagement scores of the top mainstream news and top fake news stories are subsequently compared. In the first study of this kind and perhaps the beginning of what could be called the 'fake news crisis' for Facebook, it was found that the fake news stories outperformed by engagement scores those from the mainstream news in the three-month period before the US presidential elections, thereby leading to conclusions about the comparable "power of fake election news on Facebook" (see Figure 3) (Silverman, 2016). Follow-up reporting has considered the extent to which such news continues to resonate more on Facebook than mainstream news stories, despite incipient efforts by the company to curtail its impact. One of the major studies commissioned by the US Congress found that fake news and influence campaigning activity on Facebook and especially Instagram substantially increased after the US elections (Howard et al., 2018).

In April of 2019, some two and one-half years after BuzzFeed News story, we found that only 4 of the 13 top-performing fake news and hyperpartisan websites are still online: World News Daily Report,

---

<sup>6</sup> BuzzFeed's list contains the following mainstream sources: New York Times, Washington Post, NBC News, USA Today, Politico, CNN, Wall Street Journal, CBS News, ABC News, New York Daily News, New York Post, BuzzFeed, Los Angeles Times, NPR, The Guardian, Vox, Business Insider, Huffington Post and Fox News (Silverman, 2016).

Burrard Street Journal, Twitchy and Breitbart. The others appeared to have been fly-by-night operations, which is another means of considering a source's dubiousness. That is, the other 9 sites, including two Macedonian-made ones (Denver Guardian and World Politicus) and the highest-performing site (Ending the Fed) that spread the 'Pope endorses Trump' story are all gone.

### **Facebook's adjustments**

After the US elections in 2016, Facebook CEO Mark Zuckerberg initially argued that "the idea that fake news on Facebook influenced the election in any way, I think is a pretty crazy idea," and put forward that such material amounted to a small fraction of its platform's content (Isaac, 2016). Two years later Facebook's work to quell fake news, together with its more stringent policies on (research) data access should be considered here. Addressing the fake news crisis, there has been an increase in those hired to perform 'content moderation', referred to as the janitors of social media, or even those doing the platform's dirty work (Roberts, 2017). Facebook also installed a political ad transparency tool; it lists on the ad itself who has sponsored it, and there also is a political ad archive and an API (Hern and Waterson, 2018). No longer is the targeted individual the only one able to view the hitherto 'dark post'.

With respect to algorithmic changes, in 2018 Facebook began a three-pronged strategy that would favour "meaningful connections" (family and friends), "trusted sources" (user-surveyed media) and "local news" in the news feed over more far-flung "businesses, brands, and media" (Abbruzzese, 2018; Flynn, 2018; Gartenberg, 2018). It should be remarked that these are global initiatives, coming on the heels of well-reported Facebook-associated riots in Myanmar and Sri Lanka but also the compilation of compendiums on the effects of fake news on Facebook all over the world, as the OII's global study have shown, but also the numerous governmental and think tank (umbrella) initiatives such as disinfoportal.org.

Whether Facebook's measures are working in some sense is unclear. The political ad library tool may show a source, but who is behind it may remain unclear as in the case of a pro-Brexit campaign group, Britain's Future (see Figure 4), that spent hundreds of thousands of pounds on ads in the run-up to significant UK parliamentary votes (Waterson and Hern, 2019). Efforts by journalists to unmask the source behind Britain's Future as well as other 'dark money' campaigners had for months been in vain (Monbiot, 2019). Significant political ads are also not in the archive, as ProPublica found, before its tool crowdsourcing Facebook ads and targeted individuals was purposively rendered inoperable by the company in what it called a "routine update" that would prevent illegitimate "scraping" (Merrill and Tobin, 2019). Similar tools by Mozilla and Who Targets Me also broke, thus making the verification work a difficult prospect.

The news feed tweak to boost "meaningful connections" was initially critiqued for its capacity to exaggerate the importance of fake news, as was observed in Slovakia and elsewhere when dubious sources saw their engagement scores rise (Frenkel et al., 2018). The prominence of "meaningful connections" and "local news" in the news feed, according to BuzzFeed, stirred as well as amplified the Gilets Jaunes protests in France, for their coverage on the local news made the anger groups (groupes colère) and their posts more prominent in the news feeds, as evidenced by engagement scores from Crowdtangle (Broderick and Darmanin, 2018).



Given the fake news crisis stemmed from the US elections, Facebook also has created specific initiatives for future elections that would put political parties and their positions on issues in a single, curated Facebook portal. One of the early projects was for Sweden's national elections in 2018, which, it was found in a separate study (with Twitter data), suffered from 'junk news' quantities second in magnitude only to that surrounding the US elections, and much larger in fact than such materials around the German, French and Dutch elections in 2017 (Hedman et al., 2018; Kist and Zantingh, 2017). The Facebook elections project, rolled out in meetings with social media researchers in 2018, also coincided with their new academic 'partnership' project, Social Science One. It makes available to researchers data sets such as all the URLs that have been posted to Facebook over the course of a year (King and Persily, 2018). At the same time, however, Facebook revoked approvals to research software (such as Netvizz and Netlytic) that made use of its Pages API, sparking academic protest about 'locked platforms' (Brunns et al., 2018; Rieder, 2018). Seen as reactions to the Cambridge Analytica scandal, Facebook's measures could be described as curating the datasets researchers can use. The datasets notably do not include Facebook pages themselves and their engagement scores -- data that led to the very knowledge about the fake news crisis and the scope of the Russian influence campaign in the first instance (Albright, 2017).

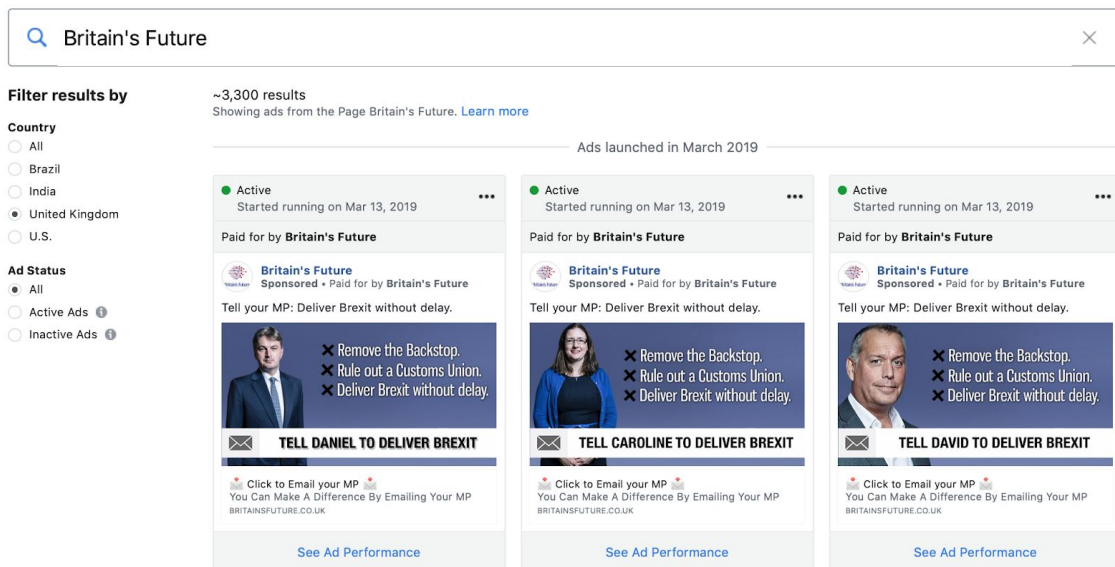


Figure 4. Facebook political ad library tool, results for Britain's Future, 13 March 2019.

### Conclusions: Fakeness and campaigning

The fake news phenomenon could be viewed as a revival of previous ones that typically have occurred when a new media technology is introduced that destabilises production, distribution and consumption of news and information, as was the case with eighteenth and nineteenth century broadsheets and tabloids (respectively) but also the radio and newswire of the twentieth century. The early web and the blogosphere also challenged existing news publication practices and were considered unedited spaces populated by self-styled authors, providing speedy news 'too fresh to be true'. Now social media platforms disrupt the trustworthiness of established news and fact, and reintroduce the idea of the web as 'truthless medium' (Marres, 2018).

The post-truth age, or condition, as it were, may be viewed in light of a conflict between what counts as 'fake' (on a source or a story level), but it has been described rather as a contest between facticity

regimes, or even sets of sectarian expertise. Locating a network of so-called 'fake news' websites, for example, could be viewed as the discovery of an influence campaign, but it just as well can be seen as an 'alternative facts' media ecology. When it is a hyperpartisan, right-wing news ecology, as in the US in the run-up to the presidential elections of 2016, it could be described as a part of the contemporary post-truth situation, or, as been often related, a culture war.

Having the 'trappings of news' in terms of look and feel, fake news has been defined as consisting of distinctive types with varying intentionality. For instance, disinformation and mal-information (the neologism) are meant to harm, whereas misinformation may be just as false but its circulation unintentional. As a case in point, satirical stories and parody may become misinformation, such as the story about the Pope endorsing Trump, which outperformed (by engagement score) any other 'news' on Facebook during the US presidential election campaigning in 2016.

In both the public as well as scholarly discourses, there has been a swing from the hype of the fake news problem (perhaps well exemplified by the Dutch tabloid headline "landslide of fake news") to its gradual debunking, e.g., "researchers say fears about 'fake news' are exaggerated" (Ingram, 2019). Such a view has resulted from a series of studies not just on engagement but also on its consumption, including the rationale behind its sharing. Small, older populations appear particularly active, as do 'heavy news consumers' and 'available audiences', or those who have at their disposal time for fringe news consumption and spreading it among online friends. The vast majority of news consumption remains of the mainstream sources, however. The evidence that consumers have been influenced or persuaded is minimal.

Nevertheless, there appears to be agreement that social media platforms remain worthy of study not only as the new 'truthless medium' but for their capacity to accelerate (or 'supercharge') fake news distribution in a hybrid media system comprised of new and established media and media formats. Despite increased content moderation, automated detection work, and a reorientation of its news feed principles, does Facebook remain a fake news machine, comparable to the one during the US presidential campaigning? Indeed, Facebook, at first hesitant to admit an issue, has taken a series of measures since then that strive to produce more trustworthiness, such as boosting posts by friends and family, crowdsourcing trusted sources as well as favouring local news, though the effectiveness of these reengineered principles has been questioned. Indeed, continuing empirical research on the most engaged-with, political news on Facebook could shed light on the quality of the platform's content delivery, however much data access may be restricted to researchers. It remains to be seen how 'oversight' research will be affected when Facebook closes research APIs and instead curates data sets for researchers, rather than allowing them to create their own. Other oversight projects have been thwarted; in early 2019 Facebook's "routine update" blocked the software by ProPublica, Mozilla and Who Targets Me that was collecting political ads and their targets, as mentioned.

The question of fake news as a campaign strategy -- be it by Russian operatives, Russified domestic actors, hyperpartisan or tendentious right-wing media-makers, and others -- also has been meticulously studied, with detailed 'playbooks' laid bare as tactics to create both a media mirage (where fact and fiction are difficult to disentangle) as well as competing truth regimes, offering counter-expertise as well as uncertainty. Governments around the world have commissioned studies, revealing the breadth and scope of the problem, explaining the playbook and putting

forward policy recommendations such as increased media literacy and the regulation of political advertising on platforms, including 'dark' posts. Platforms are asked to create public archives, which also would benefit research as well as (data) journalism. Fact-checking also has gone global, though it often remains a small-scale enterprise practiced by bespoke bureaus, occasionally working in tandem with Facebook, checking posts that have been flagged by users, and weighing in on the question of fakery.

Finally, there are scholars in the US and recently in Europe putting forward the argument that studying Russian disinformation shifts the attention away from the home-grown hyperpartisan news ecologies that have been emerging over the past few years, particularly on the right (Benkler et al., 2017; Benkler et al., 2018; Rone, 2019). The point also fits with the 'dark globalization' argument concerning how existing domestic divisions, displayed in this media, may be exacerbated by foreign operatives but are not created by them. To date the effectiveness of Russian influence campaigning in Europe, in either sowing or exacerbating division, has yet to be compellingly demonstrated; the false and junk domestic news sources (e.g., the pro-Russian sources re-narrating the cause of the downing of MH17) also appear to have scant reach (Fletcher et al., 2018). In a climate of heightened sensitivity towards dubious sources and stories, it remains to be seen whether they have staying power.

## References

Abbruzzese, Jason (2018) 'Facebook became your news diet. Now, it's going to serve you junk', *Mashable*, 17 January.

Access Now, Civil Liberties Union For Europe, and European Digital Rights (2018) *Informing the "disinformation" debate*, Report, Published online on October 18.

Albright, Jonathan (2017) 'Itemized Posts and Historical Engagement - 6 Now-Closed FB Pages', data set, Tableau Public, 5 October, <https://public.tableau.com/profile/d1gi#!/vizhome/FB4/TotalReachbyPage>.

Allcott, Hunt and Matthew Gentzkow (2017) 'Social Media and Fake News in the 2016 Election', *Journal of Economic Perspectives*, 31(2): 211–236.

Alvarez, R. Michael, Ines Levin, Alexander H. Trechsel and Kristjan Vassil (2014) 'Voting Advice Applications: How Useful and for Whom?', *Journal of Information Technology & Politics*, 11(1): 82-101, DOI: 10.1080/19331681.2013.873361

Angwin, Julia, Madeleine Varner and Ariana Tobin (2017) 'Facebook Enabled Advertisers to Reach "Jew Haters"', *ProPublica*, 14 September.

Barbrook, Richard (1998) 'The Hi-Tech Gift Economy', *First Monday*, 3(12).

Beiler, Markus and Johanna Kiesler (2018) "'Lügenpresse! Lying press!" Is the Press Lying?', in Otto, Kim and Andreas Köhler (eds.) *Trust in Media and Journalism*, Wiesbaden: Springer, 155-179.

Belk, Russell (2007) 'Why Not Share Rather Than Own?', *The ANNALS of the American Academy of Political and Social Science*, 611: 126-140.

Bell, Emily and Taylor Owen (2017) 'The Platform Press', Report, Columbia University: Tow Center for Journalism.

Bellingcat Investigation Team (2016) 'Behind the Dutch Terror Threat Video: The St. Petersburg "Troll Factory" Connection', *Bellingcat*, 3 April.

Benkler, Yochai, Robert Faris, Hal Roberts, and Ethan Zuckerman (2017) 'Study: Breitbart-Led Right-Wing Media Ecosystem Altered Broader Media Agenda', *Columbia Journalism Review*, 3 March.

Benkler, Yochai, Robert Faris and Hal Roberts (2018) *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford: Oxford University Press.

Bessi, Alessandro and Emilio Ferrara (2016) 'Social bots distort the 2016 US presidential election online discussion', *First Monday*, 21(11).

Boczkowski, Pablo (2016) 'Fake news and the future of journalism', *NiemanLab*, December.

Boghardt, Thomas (2009) 'Soviet Bloc Intelligence and Its AIDS Disinformation Campaign', *Studies in Intelligence*, 53(4): 1-24.

Bounegru, Liliana, Jonathan Gray, Tommaso Venturini, Michele Mauri (2018) *A Field Guide to "Fake News" and Other Information Disorders: A Collection of Recipes for Those Who Love to Cook with Digital Methods*. Amsterdam: Public Data Lab.

Bradshaw, Samantha and Phillip N. Howard (2018) 'Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation', Computational Propaganda Data Memo, Oxford: Oxford Internet Institute.

Bradshaw, Samantha (2018) 'Responding to Fake News through regulation and automation', *Fake News, Authentic Views*, London: Carter-Ruck.

Brattberg, Erik and Tim Mauer (2018) 'Russian Election Interference: Europe's Counter to Fake News and Cyber Attacks', White Paper, Washington, DC: Carnegie Endowment for International Peace, May.

Broderick, Ryan and Jules Darmanin (2018) 'The "Yellow Vest" Riots In France Are What Happens When Facebook Gets Involved With Local News', *Buzzfeed News*, 6 December.

Brennen, Bonnie (2017) 'Making Sense of Lies, Deceptive Propaganda, and Fake News', *Journal of Media Ethics*, 32(3): 179-181.

Bruns, Axel, Anja Bechmann, Jean Burgess et al. (2018) 'Facebook shuts the gate after the horse has bolted, and hurts real research in the process', *Internet Policy Review*, 25 April.

Bump, Philip (2017) 'The investigation goes digital: Did someone point Russia to specific online targets?' *Washington Post*, 12 July.

Burger, Peter (2016) 'Moslims vernielen kerstmarkt in Litouwen? Nee: rellen in de VS', *De Gestolen Grootmoeder blog*, 14 December.

Burger, Peter, Soeradj Kanhai, Alexander Pleijter, Suzan Verberne (2019) 'The reach of commercially motivated junk news on Facebook', *arXiv* pre-print, forthcoming.

Butac, Maicolengel (2018) 'The Reuters Institute for the Study of Journalism vs fake news', *BUTAC blog*, 8 February.

Cadwalladr, Carole and Emma Graham-Harrison (2018) 'Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach', *The Guardian*, 17 March.

Canadian Security Intelligence Service (2018) *Who said what? The Security Challenges of Modern Disinformation*, World Watch: Expert Notes series publication (No. 2016-12-05), Canada: Canadian Security Intelligence Service.

Cary, Stephen G. (1955) *Speak Truth to Power: A Quaker Search for an Alternative to Violence*, Philadelphia: American Friends Service Committee.

Chadwick, Andrew (2013) *The Hybrid Media System: Politics and Power*. Oxford: Oxford University Press.

Chen, Adrian (2015) 'The Agency', *The New York Times*, 2 June.

Claussen, Victor (2018) 'Fighting hate speech and fake news. The Network Enforcement Act (NetzDG) in Germany in the context of European legislation', *Media Laws 3*, published online on October 14.

Colbert Report (2005) 'The WØRD "Truthiness"', *Comedy Central*, 17 October.

Collins, Keith Sheera Frenkel (2018). 'Can you spot the deceptive Facebook post?' *New York Times*, 4 September.

Dahlgren, Kristen and Daniel Arkin (2017) '11-Year-Old Texas Boy Invents Device to Prevent Hot Car Deaths', *NBC News*, 29 June.

Daniels, Jessie (2018) 'The algorithmic rise of the alt-right', *Contexts*, 17(1): 60-65.

Darnton, Robert (2010) *Poetry and the Police: Communication Networks in Eighteenth-Century Paris*. Cambridge, MA: Harvard University Press.

Darnton, Robert (2017) 'The True History of Fake News', *New York Review of Books*, 13 February.

Day, Amber and Ethan Thompson (2012) 'Live from New York, it's the fake news! Saturday night live and the (non)politics of parody,' *Popular Communication* 10(1–2): 170–182.

DeAmicis, Carmel (2014) 'Facebook shifts its algorithm to fight clickbait. Will it kill off Upworthy and BuzzFeed?', *GIGAOM*, Austin: Giga Omni Media.

Drog (2018) 'Slecht Nieuws: Serious game over propaganda', website, <https://www.slechtnieuws.nl/>.

Duke Reporters Lab (2019) *Fact Checking News*, website, <https://reporterslab.org/fact-checking/>

Dumont, P. and Kies, R. (2012) 'Smartvote.lu: Usage and impact of the first VAA in Luxembourg', *International Journal of Electronic Governance* 5(3/4): 388–410.

Ellick, Adam B. and Adam Westbrook (2018) 'Operation Infektion', *New York Times*, 12 November.

EU HLEG (2018) *A multi-dimensional approach to disinformation*, Report of the independent High level Group on fake news and online disinformation, Luxembourg: Publications Office of the European Union.

Fandos, Nicholas (2017) 'White House Pushes 'Alternative Facts.' Here Are the Real Ones', *New York Times*, 22 January.

Faris, Robert M., Hal Roberts, Bruce Etling, Nikki Bourassa, Ethan Zuckerman, and Yochai Benkler (2017) 'Partisanship, Propaganda, and Disinformation: Online Media and the 2016 U.S. Presidential Election', Berkman Klein Center for Internet & Society Research Paper.

Farkas, Johan and Marco Bastos (2018) 'IRA Propaganda on Twitter: Stoking Antagonism and Tweeting Local News', *SMSociety '18*, Copenhagen.

FireEye (2018) 'Suspected Iranian Influence Operation', Report, Milpitas, CA: FireEye.

Fitts, Alexis Sobel (2017) 'Welcome to the Wikipedia of the Alt-right', *Wired*, 21 June.

Fletcher, Richard, Alessio Cornia, Lucas Graves, and Rasmus Kleis Nielsen (2018) 'Measuring the reach of "fake news" and online disinformation in Europe', Reuters Institute for the Study of Journalism, Oxford University, February.

Flynn, Kerry (2018) 'Facebook will elevate "trusted" news outlets after surveying U.S. users', *Mashable*, 19 January.

Frenkel, Sheera, Kate Conger and Kevin Roose (2019) 'Russia's Playbook for Social Media Disinformation Has Gone Global', *New York Times*, 31 January.

Frenkel, Sheera, Nicholas Casey and Paul Mozur (2018) 'In Some Countries, Facebook's Fiddling Has Magnified Fake News', *New York Times*, 14 January.

Fubini, Federico (2018) 'Tweet populist dalla Russia sulla politica italiana. Come negli Usa', *Corriere della Sera*, 1 August.

Gartenberg, Chaim (2018) 'Facebook is going to start surveying users to determine trustworthy news sources', *The Verge*, 19 January.

Garzia, Diego and Marschall, Stefan (2012) 'Voting Advice Applications Under Review: The State of Research', *International Journal of Electronic Governance*, 5: 203-22.

German Federal Foreign Office (2018) 'Fake news, bots and provocative statements – disinformation on the internet', *Auswaertiges Amt website*, published online on August 7.

Gessen, Masha (2018) 'Why the Russian Influence Campaign Remains So Hard to Understand', *The New Yorker*, 18 December.

Gitelman, Lisa (2006) *Always Already New: Media, History, and the Data of Culture*. Cambridge, MA: MIT Press.

Gottfried, Jeffrey and Elisa Shearer (2016) 'News Use Across Social Media Platforms 2016', Washington, DC: Pew Research Center.

Government Offices of Sweden (2017) *A practical approach on how to cope with disinformation*, Government Offices of Sweden website, published online on October 6.

Gramer, Robbie (2017) 'Denmark Creates the World's First Ever Digital Ambassador', *Foreign Policy*, published online on January 27.

Graves, Lucas (2016) *Deciding What's True: The Rise of Political Fact-Checking in American Journalism*. New York: Columbia University Press.

Graves, Lucas and Federica Cherubini (2016) *The Rise of Fact-checking Sites in Europe*. Oxford: University of Oxford, Reuters Institute for the Study of Journalism.

Grinberg, Nir, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson and David Lazer (2019) 'Fake news on Twitter during the 2016 U.S. presidential election', *Science* 363(6425): 374-378.

Groll, Elias (2018) 'Battling the Bots', *Foreign Policy*, 12 November.

Guess, Andrew, Brendan Nyhan and Jason Reifler (2018) 'Selective Exposure to Misinformation: Evidence from the consumption of fake news during the 2016 U.S. presidential campaign', Report, Hannover, NH: Dartmouth College, <https://www.dartmouth.edu/~nyhan/fake-news-2016.pdf>.

Haciyakupoglu, Gulizar, Jennifer Yang Hui, V. S. Suguna, Dymples Leong, and Muhammad Faizal Bin Abdul Rahman (2018) *Countering Fake News: A Survey Of Recent Global Initiatives*, Singapore: Nanyang Technological University.

Hall, Jim (2001) *Online journalism: a critical primer*. London: Pluto Press.

Hameleers, Michael and Toni van der Meer (2019) 'Fact-checks helpen tegen desinformatie! Maar dan moeten ze wel gezien worden', *De Nieuwe Reporter*, 29 January.

Harvey, Del and Yoel Roth (2018) 'An update on our elections integrity work', Twitter blog, 1 October.

Haynes, Gavin (2016). 'Can you spot the 'real' fake news story?' *The Guardian*, 28 December.

Heck, Wilmer (2018). 'Nederlandse media dagen EU voor rechter na beschuldigingen desinformatie', *NRC Handelsblad*, 20 February.

Hedman, Freja, Fabian Sivnert, Lisa-Maria Neudert, Bence Kollanyi, Philip N. Howard and Vidya Narayanan (2018) 'News and Political Information Consumption in Sweden: Mapping the 2018 Swedish General Election on Twitter', Computational Propaganda Data Memo, Oxford: Oxford Internet Institute, 6 September.

Hedrick, Ashley, Dave Karpf and Daniel Kreiss (2018) 'The Earnest Internet vs. the Ambivalent Internet', *International Journal of Communication*, 12: 1057-1064.

Hendricks, Vincent F. and Mads Vestergaard (2019) *Reality Lost: Markets of Attention, Misinformation and Manipulation*. Cham: Springer.

Hern, Alex (2017) 'Google acts against fake news on search engine', *The Guardian*, 25 April.

Hern, Alex and Jim Waterson (2018) 'Facebook cracks down on 'dark ads' by British political groups', *The Guardian*, 16 October.

Herrman, John (2016) 'Inside Facebook's (Totally Insane, Unintentionally Gigantic, Hyperpartisan) Political-Media Machine', *The New York Times*, 28 August.

Hirzalla, Fadi and Liesbet van Zoonen (2015) 'Online voting applications. Foci, findings and future of an emerging research field' in Coleman, S. & Freelon, D. (eds), *Handbook of Digital Politics*, Northampton, MA: Edward Elgar: 87-103.

Holzman, Michael (2008) *James Jesus Angleton, the CIA and the Craft of Counterintelligence*, Amherst: University of Massachusetts Press.

House of Commons (2019) 'Disinformation and "Fake News": Final Report Eighth Report of Session 2017–19', Digital, Culture, Media and Sport Committee, London: UK Parliament.



Howard, Philip N., Gillian Bolsover, Bence Kollanyi, Samantha Bradshaw, and Lisa-Maria Neudert (2017) 'Junk news and bots during the U.S. election: What were michigan voters sharing over Twitter?' Computational Propaganda Data Memo, Oxford: Oxford Internet Institute.

Howard, Philip N., Bharath Ganesh, Dimitra Liotsiou, John Kelly and Camille François (2018) 'The IRA, Social Media and Political Polarization in the United States, 2012-2018' Report, Computational Propaganda Research Project, Oxford: Oxford Internet Institute.

Isaac, Mike (2016) 'Facebook, in Cross Hairs After Election, Is Said to Question Its Influence', *New York Times*, 12 November.

Jack, Caroline (2017) *Lexicon of Lies: Terms for Problematic Information* New York: Data & Society Research Institute.

Jankowski, Nicholas W. (2018) 'Researching Fake News: A Selective Examination of Empirical Studies', *Javnost - The Public*, 25(1-2): 248-255.

Jeangène Vilmer, Jean-Baptiste, Alexandre Escorcica, Marine Guillaume, Janaina Herrera (2018) *Information Manipulation: A Challenge for Our Democracies*, report by the Policy Planning Staff (CAPS) of the Ministry for Europe and Foreign Affairs and the Institute for Strategic Research (IRSEM) of the Ministry for the Armed Forces, Paris, August.

Jonker, Jorn (2017) 'Nepnieuws gevaar voor Nederland', *De Telegraaf*, 14 November.

Kanne, Peter and Milan Driessen (2017) *Desinformatie leidt tot verwarring bij nieuwsconsument*, Amsterdam: I&O Research.

Kasteleijn, Nando (2017) 'Nepnieuws groot gevaar voor Nederland? Dat lijkt dus wel mee te vallen', *NOS*, 28 August.

van Keulen, Ira, Iris Korthagen, Paul Diederer en Pieter van Boheemen (2018) 'Digitalisering van het nieuws: Online nieuwsgedrag, desinformatie en personalisatie in Nederland', Den Haag: Rathenau Instituut.

King, Gary and Nathaniel Persily (2018) 'A New Model for Industry-Academic Partnerships', Working Paper, 9 April, <http://j.mp/2q1IQpH>.

Kist, Reinier (2019) 'Factchecken Facebook loopt stuk op aansprakelijkheid', *NRC Handelsblad*, 26 February.

Kist, Reinier and Rik Wassens (2017) 'Ook Nederlandse media werden misleid door Russische trollen', *NRC Handelsblad*, 8 December.

Kist, Reinier and Rik Wassens (2018a) 'Russische trollen actief in Nederland', *NRC Handelsblad*, 15 July.

- Kist, Reinier and Rik Wassens (2018b) 'Russisch trollenleger ook actief in Nederland', *NRC Handelsblad*, 15 July.
- Kist, Reinier and Peter Zantingh (2017) 'Geen grote rol nepnieuws in aanloop naar verkiezingen', *NRC Handelsblad*, 6 March.
- Kitta, Andrea (2018) 'Alternative Websites and Fake News: Taking a Stab at Definition, Genre and Belief', *The Journal of American Folklore*, 131(522): 405-412.
- Kouwenhoven, Andreas and Hugo Logtenberg (2017) 'Hoe Denk met 'trollen' politieke tegenstanders monddood probeert te maken', *NRC Handelsblad*, 10 February.
- Kranenberg, Annieke (2017) 'Wie weet nog wat er waar is?', *de Volkskrant*, 23 December.
- Latour, Bruno (2008) *What is the style of matters of concern?* Assen: Van Gorcum.
- Laquintano, Timothy and Annette Vee (2017) 'How Automated Writing Systems Affect the Circulation of Political Information Online', *Literacy in Composition Studies*, 5(2): 43-62.
- Lazarsfeld, Paul F., Bernard R. Berelson and Hazel Gaudet (1948) *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign*. New York: Columbia University Press.
- Lazer, David M.J., Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A. Sloman, Cass R. Sunstein, Emily A. Thorson, Duncan J. Watts and Jonathan L. Zittrain (2018) 'The science of fake news', *Science*, 359(6380): 1094-1096.
- Lippmann, Walter (1922) *Public Opinion*, New York: Harcourt, Brace and Co.
- Louwerse, Tom and Martin Rosema (2014) 'The design effects of voting applications: Comparing methods of calculating matches', *Acta Politica* 49(3):286-312.
- Lüders, Marika, Lin Prøitz and Terje Rasmussen (2010) 'Emerging personal media genres', *New Media & Society*, 12(6) 947-963.
- Margolis, Joseph (1995). 'Beyond postmodernism: Logic as rhetoric', *Argumentation*, 9(1): 21-31.
- Marres, Noortje (2018) 'Why We Can't Have Our Facts Back', *Engaging Science, Technology, and Society* 4, 423-443.
- Marvin, Carolyn (1988) *When Old Technologies were New*. New York: Oxford University Press.
- Marwick, Alice E. (2018) 'Why Do People Share Fake News? A Sociotechnical Model of Media Effects', *Georgetown Law Technology Review*, 2(2): 474-512.

McGonagle, Tarlach (2017). "'Fake news': False fears or real concerns?", *Netherlands Quarterly of Human Rights*, 35(4): 203–209.

McKernon, Edward (1925) 'Fake News and the Public', *The Harper's Monthly* October, 528-536.

McKernon, Edward (1928) 'News Fakers', *The Outlook* 149(4): 130-141.

McNeil, Maureen (2013) 'Between a Rock and a Hard Place: The Deficit Model, the Diffusion Model and Publics in STS', *Science as Culture*, 22(4): 589-608.

McQueen, Sharon (2018) 'From Yellow Journalism to Tabloids to Clickbait: The Origins of Fake News in the United States', in Agosto, Denise E. (ed.) *Information Literacies and Libraries in the Age of Fake News*, Santa Barbara, CA: Libraries Unlimited, 12-35.

Merrill, Jeremy B. and Ariana Tobin (2019) 'Facebook Moves to Block Ad Transparency Tools — Including Ours', *ProPublica*, 29 January.

Milneil, Christian (2018) Data: Read the tweets from alleged Russian troll accounts targeting Maine politicians', *Portland Press Herald*, 2 August.

Mina, An Xiao (2019). *Memes to Movements: How the World's Most Viral Media Is Changing Social Protest and Power*, Boston: Beacon Press.

Misérus, Mark and Robert van der Noordaa (2018a) 'Het trollenleger van popartiest Dotan', *de Volkskrant*, 14 April.

Misérus, Mark and Robert van der Noordaa (2018b) 'Dotan erkent aanmaken nepfans: "Ik was heel naïef, veel te ambitieus en onzeker"', *de Volkskrant*, 16 April.

Monbiot, George (2019) 'Dark money is pushing for a no-deal Brexit. Who is behind it?' *The Guardian*, 13 February.

Morozov, Evgeny (2017) 'Moral panic over fake news hides the real enemy – the digital giants', *The Guardian*, 8 January.

Neudert, Lisa-Maria N. (2017) 'Computational Propaganda in Germany: A Cautionary Tale', Computational Propaganda Working Paper 2017.7, Oxford: Oxford Internet Institute.

New Knowledge (2018) 'The Tactics & Tropes of the Internet Research Agency', White Paper, Austin, TX: New Knowledge.

Newman, Michelle C. (2010) 'The Daily Show and Meta-Coverage: How Mock News Covers the Political Communications System', *The Elon Journal of Undergraduate Research in Communications*, 1(2).

NewsGuard (2019), 'Restoring Trust and Accountability', webpage, <https://www.newsguardtech.com>.

Nijmeijer, Bert (2018) 'Zelf nepnieuws maken om daarna de echte trollen te kunnen herkennen', *NRC Handelsblad*, 22 July.

van der Noorda, Robert and Coen van de Ven (2018a) 'Hoe Russische trollen inspelen op westerse angsten', *De Groene Amsterdammer*, 28 Augustus.

van der Noordaa, Robert and Coen van de Ven (2018b) '3 miljoen "trollentweets" onderzocht: hoe Rusland via sociale media ook in België verdeeldheid zaait', *Knack*, 27 November.

NOS (2017a) 'Baby van 20 kilo heeft de primeur: eerste nepnieuws in Nederland', *NOS*, 8 June.

NOS (2017b) 'Ollongren: Russische desinformatie bij Oekraïne-referendum', *NOS*, 15 November.

NOS (2018) 'Tweeters grote schoonmaak: Wilders en Denk-politici verliezen volgers', *NOS*, 13 July.

Nuzzi, Olivia (2017) 'Kellyanne Conway Is a Star', *New York Magazine*, March.

Nyhan, Brendan, Ethan Porter, Jason Reifler and Thomas J. Wood (2019) 'Taking Fact-Checks Literally But Not Seriously? The Effects of Journalistic Fact-Checking on Factual Beliefs and Candidate Favorability', *Political Behavior*, published online 21 January.

O'Brien, Chris (2019) 'Sheryl Sandberg says Facebook is now blocking 1 million fake accounts every day', *Venture Beat*, 21 January.

O'Donovan, Caroline (2014) 'What is clickbait?' *Niemanlab*, 25 August.

Opper, F. (1894) 'The fin de siècle newspaper proprietor', *Puck* 35(887).

Panagopoulos, Costas (2016) 'All about that base: Changing campaign strategies in US presidential elections', *Party Politics*, 22(2): 179–90.

Pariser, Eli (2011) *The Filter Bubble*. New York: Penguin.

Parlapiano, Alicia and Jasmine C. Lee (2018) 'The Propaganda Tools Used by Russians to Influence the 2016 Election', *New York Times*, 16 February.

Peck, Reece (2019) *Fox Populism: Branding Conservatism as Working Class*. Cambridge: Cambridge University Press.

Pedersen, David B. and Vincent F. Hendricks (2014) 'Science Bubbles', *Philosophy and Technology*, 27(4): 503-518.

Persily, Nathaniel (2017) 'The 2016 U.S. Election: Can Democracy Survive the Internet?' *Journal of Democracy*, 28(2): 63-76

Phillips, Whitney (2015) *This Is Why We Can't Have Nice Things: Mapping the Relationship Between Online Trolling and Mainstream Culture*. Cambridge, MA: MIT Press.

Phillips, Whitney (2018) 'The Oxygen of Amplification', Report, New York: Data & Society Research Institute.

Phillips, Whitney and Ryan Milner (2018) *The Ambivalent Internet: Mischief, Oddity, and Antagonism Online*, Cambridge: Polity.

Pleijter, Alexander (2017) 'De nepnieuwslawine zonder nepnieuws', *Villamedia*, 17 November.

Pohjonen, Matti and Sahana Udupa (2017) 'Extreme Speech Online: An Anthropological Critique of Hate Speech Debates', *International Journal of Communication* 11: 1173-1191.

Posetti, Julie and Alice Matthews (2018) *A short guide to the history of 'fake news' and disinformation. A learning module for journalists and journalism educator*. Washington, DC: International Center for Journalists.

Poynter (2019) 'A guide to anti-misinformation actions around the world', St. Petersburg, FL: The Poynter Institute.

Puschmann, Cornelius (2018) 'Beyond the Bubble: Assessing the Diversity of Political Search Results', *Digital Journalism*, published online 28 November.

Ratkiewicz, Jacob, Michael D. Conover, Mark Meiss, Bruno Goncalves, Alessandro Flammini, Filippo Menczer Menczer (2011) 'Detecting and Tracking Political Abuse in Social Media', *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, Barcelona, Spain, July.

Reid, John, Bastiaan Geleijnse and Jean-Marc van Tol (2018) 'Fokke en Sukke hebben zitting in de taskforce nepnieuws', Cartoon, 17 January.

Renner, Nausicaa (2017) 'Memes trump articles on Breitbart's Facebook page', *Columbia Journalism Review*, 30 January.

RoBhat Labs (2017) 'Identifying Propaganda Bots on Twitter', *Medium*, 31 October.

Rieder, Bernhard (2018) 'Facebook's app review and how independent research just got a lot harder', Politics of Systems blog, 11 August.

Rieder, Bernhard, Ariadna Matamoros-Fernández and Òscar Coromina (2018). "From ranking algorithms to 'ranking cultures': Investigating the modulation of visibility in YouTube search results," *Convergence: The International Journal of Research into New Media Technologies*, 24(1): 50–68.

RNTC (2019). 'Challenging fake news', webpage, Hilversum: RNTC, <https://rntc.com/blog/challenging-fake-news>.

Roberts, Sarah T. (2016) 'Commercial Content Moderation: Digital Laborers' Dirty Work,' in Safiya Umoja Noble and Brendesha M. Tynes (eds.) *The Intersectional Internet*. New York: Peter Lang, 147-160.

Roeder, Oliver (2018) 'We Gave You 3 Million Russian Troll Tweets. Here's What You've Found So Far', *FiveThirtyEight*, 8 August.

Rogers, Richard (2005) 'Poignancy in the US political blogosphere', *Aslib Proceedings: New Information Perspectives*, 57(4): 356-368.

Rogers, Richard (2018a) 'Otherwise Engaged: Social Media from Vanity Metrics to Critical Analytics', *International Journal of Communication*, 12: 450-472.

Rogers, Richard (2018b). Digital Methods for Cross-Platform Analysis, in Burgess, Jean, Alice Marwick and Thomas Poell (eds.), *SAGE Handbook of Social Media*. London: Sage, 91-110.

Rone, Julia (2019) 'Why talking about "disinformation" misses the point when considering radical right "alternative" media', LSE Media Policy project blog, 9 January.

Rony, Md Main Uddin, Naeemul Hassan and Mohammad Yousuf (2017) 'BaitBuster: Destined to Save You Some Clicks', *Proceedings of Computation+Journalism Symposium*, Northwestern University, October.

Roozenbeek, Jon and Sander van der Linden (2018) 'The fake news game: actively inoculating against the risk of misinformation', *Journal of Risk Research*, DOI: 10.1080/13669877.2018.1443491.

Ruusuvirta, Outi (2010) 'Much Ado About Nothing? Online Voting Advice Applications in Finland', in Cedroni, Lorella and Diego Garzia (eds.), *Voting Advice Applications in Europe. The State of the Art*, Napoli, ScriptaWeb: 47-77.

Sängerlaub, Alexander, Miriam Meier and Wolf-Dieter Rühl (2017) 'Fakten statt Fakes: Das Phänomen "Fake News"', Berlin: Stiftung Neue Verantwortung.

Shane, Scott (2018) 'How Unwitting Americans Encountered Russian Operatives Online', *New York Times*, 18 February.

Shao, Chengcheng, Giovanni Luca Ciampaglia, Onur Varol, Kai-Cheng Yang, Alessandro Flammini and Filippo Menczer (2018) 'The spread of low-credibility content by social bots', *Nature Communications*, 9, 4787, 20 November.

Shaw, Eugene (1979) 'Agenda-Setting and Mass Communication Theory', *International Communication Gazette*, 25(2): 96-105.

Silverman, Craig, J. Lester Feder, Saska Cvetkovska and Aubrey Belford (2018) 'Macedonia's Pro-Trump Fake News Industry Had American Links, And Is Under Investigation For Possible Russia Ties', *Buzzfeed News*, 18 July.

Silverman, Craig (2016) 'This Analysis Shows How Viral Fake Election News Stories Outperformed Real News On Facebook', *Buzzfeed News*, 16 November.

Silverman, Craig and Lawrence Alexander (2016) 'How Teens In The Balkans Are Duping Trump Supporters With Fake News', *Buzzfeed News*, 3 November.

Silverman, Craig and Jeremy Singer-Vine, (2016) 'The True Story Behind The Biggest Fake News Hit Of The Election', *Buzzfeed News*, 16 December.

Silverman, Craig, Lauren Strapagiel, Hamza Shaban, Ellie Hall and Jeremy Singer-Vine (2016) 'Hyperpartisan Facebook Pages Are Publishing False And Misleading Information At An Alarming Rate', *Buzzfeed News*, 20 October.

Sismondo, Sergio (2017) Post-truth?, *Social Studies of Science*, 47(1): 3–6.

Snyder, Timothy (2018) *The Road to Unfreedom: Russia, Europe, America*. New York: Tim Duggan Books.

Solon, Olivia and Sam Levin (2016) 'How Google's search algorithm spreads false information with a rightwing bias', *The Guardian*, 16 December.

Srnicek, Nick (2017) *Platform Capitalism*. Cambridge: Polity.

Still, Keith (2017) 'Inauguration Crowd Size', Crowd Safety and Risk Analysis blog, Manchester Metropolitan University, <http://www.gkstill.com/CV/Projects/Counting.html>.

Swedish Civil Contingencies Agency (2018) *Countering information influence activities: A handbook for communicators*, Karlstad: Swedish Civil Contingencies Agency (MSB).

Swedish Media Council (2019) *Källkritik — en utmaning*, webpage, <https://statensmedierad.se/larommedier/kallkritikvemvadvarfor/kallkritikenutmaning.422.html>, 5 March.

Tandoc Jr., Edson C., Zheng Wei Lim and Richard Ling (2018) 'Defining "Fake News": A typology of scholarly definitions', *Digital Journalism*, 2: 137-153.

Timberg, Craig (2017) 'Russian propaganda may have been shared hundreds of millions of times, new research says', *Washington Post*, 5 October.

Tripodi, Francesca (2018). 'Alternative Facts, Alternative Truths', *Data & Society: Points*, 23 February.

Tucker, Joshua A., Andrew Guess, Pablo Barberá, Cristian Vaccari, Alexandra Siegel, Sergey Sanovich, Denis Stukal, and Brendan Nyhan (2018) 'Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature', report, Menlo Park, CA: William and Flora Hewlett Foundation.

Turner, Stephen (2001) 'What is the problem with experts?', *Social Studies of Science*, 31(1): 123–149.

Tuters, Marc, Emilija Jokubauskaitė and Daniel Bach (2018) 'Post-Truth Protest: How 4chan Cooked Up the Pizzagate Bullshit', *M/C: A Journal of Media and Culture*, 21(3).

Vaidhyanathan, Siva (2018) 'Why Facebook Will Never Be Free of Fakes', *New York Times*, 5 September.

Vaidhyanathan, Siva (2017) 'Facebook Wins, Democracy Loses', *New York Times*, 8 September.

Vassil, Kristjan (2011). *Voting smarter? The impact of voting advice applications on political behavior*, PhD dissertation, European University Institute, Tartu, Estonia.

Volpicelli, Gina (2019) 'This is how Facebook's news feed fact-checking will work in the UK', *Wired*, 11 January.

Vosoughi, Soroush, Deb Roy and Sinan Aral (2018) 'The spread of true and false news online', *Science*, 359(6380): 1146-1151.

Walgrave, Stefaan, Michiel Nuytemans and Koen Pepermans (2009) 'Voting Aid Applications and the Effect of Statement Selection', *West European Politics* 32(6): 1161-1180.

Wardle, Claire (2018) *Information Disorder: The Essential Glossary*, Shorenstein Center, Cambridge, MA: Harvard Kennedy School.

Wardle, Claire and Hossein Derakhshan (2017) *Information Disorder: Toward an interdisciplinary framework for research and policy making*, Strasbourg: Council of Europe.

Wendling, Mike (2018) "The (almost) complete history of 'fake news'", *BBC News*, 22 January.

Wieringa, Maranke, Tim de Winkel and Callum Lewis (2017) 'Wie is de waakhond op sociale media?', report, Utrecht: Utrecht Data School.

Woolley, Samuel and Philip N. Howard (2016) 'Social media, revolution, and the rise of the political Bot', in Piers Robinson, Philip Seib, Romy Fröhlich (eds.), *Handbook of Media, Conflict and Security*. New York: Routledge, 282-292.

Wynne, Brian (1991) Knowledge in context, *Science, Technology & Human Values*, 16(1): 111-121.



van Zijl, Frank and Huib Modderkolk (2017) 'Generaal: Nederland kan digitale dreiging niet aan', *De Volkskrant*, 29 August.

\* \* \* \* \*

## Appendix

### **Governmental efforts and discussions of countermeasures**

A first step for many national governments and other regional political entities that wish to counter disinformation is to install committees as well as task forces; it occurs across the globe, from the much publicised hearings by the US Congress and UK Parliament on the Russian involvement in the US elections and the Cambridge Analytica affair, to the task forces and other entities formed in many of the nearly 50 countries where influence campaigning has taken place (Bradshaw and Howard, 2018). Following from these convenings, there have been national calls to regulate the “digital giants”, and the European Union, through its creation of a High-Level Expert Group (EU HLEG) on fake news and online disinformation, has issued its recommendations for countering disinformation, including calls for transparency, media and information literacy, and tools for empowering journalism. In the European countries with recent or imminent national elections there has been even greater urgency, with Germany and France enacting legislation (online hate speech and ‘fake news laws’, respectively), and Sweden and Denmark engaging in awareness-raising as well as media literacy campaigns. Denmark installed a ‘digital ambassador’ (Gramer, 2017).

Below is a list of certain measures to counteract disinformation and fake news, gleaned from recent governmental documents and related materials. They include social media company regulation, codes of ethics, fact-checking and media literacy campaigning.

#### *Social media company regulation*

Many government committees agree that the large tech companies that have come to dominate the online realm, such as Google, Twitter, and Facebook, should be regulated, but caution over-regulation in forms that would curtail expression and press freedoms. The starting point for the regulation of these companies to counter disinformation is to address political advertising on social media platforms. It can include the verification of those paying for political advertisements and disclosing them publicly. Additionally, the social media companies could be required to create public archives of advertisements so that among other ad types ‘dark posts’ may be studied (Bradshaw, 2018). In fact, Facebook has such an archive (and an API), but it also prevented watchdogs including Mozilla from verifying its collection techniques, equating their methods with illegitimate data “scraping” (Merrill and Tobin, 2019).

Relatedly, the EU HLEG proposes the development of a ‘European-wide code of practices’ that describes the roles and responsibilities of relevant stakeholders such as tech companies, and media organizations but also research organizations and fact-checking initiatives, based on key principles (2018). In short, they address the adaptation of political advertising policies (including sponsored advertisements and other forms of content), and the provision of access to data for research and fact-checking. They also propose the installation of advanced settings for users to customise their user experience, collaboration with news outlets to facilitate users’ access to trustworthy news, the facilitation of fact-checking and content flagging, and allowing users to “exercise their right to reply” (EU HLEG, 2018: 32-33).

The UK Parliamentary report on fake news and disinformation speaks in an unusually piqued tone of the importance of regulating social media platforms and related tech companies, singling out Facebook as providing the “impression of working towards transparency”, but often “obfuscating”

how well it is capturing and archiving political ads (House of Commons, 2019: 85). Ultimately they call for establishing an “educational levy” or charge on social media companies to fund digital literacy as a fourth pillar of the education system after reading, writing and maths (House of Commons, 2019: 87). There is also a recommendation that social media companies should develop means to distinguish between those sources regularly furnishing disinformation and those who do not, in a new system of “content regulation” (House of Commons, 2019: 87). While carefully worded, that measures can count on the criticism that similar proposals have faced concerning the restriction of the freedom of expression, while not being effective measures against hateful or incendiary content (Access Now et al., 2018).

Nevertheless, legislation has been passed. Germany has established a law, NetzDG, that extends its hate speech legislation compelling social media companies (with more than two million registered users in Germany) to remove such speech within twenty-four hours or face hefty fines (Claussen, 2018). More controversially, France has new legislation which applies to ‘false information’; the law requires that three months prior to an election ‘false news’ be removed.

#### *Detecting and removing fake content*

The Reporters’ Lab at Duke University keeps track of fact-checking initiatives worldwide, and has identified some 160 active initiatives (Duke Reporters Lab, 2019). In European countries, some fact-checking initiatives are attached to news organisations, but most are operating as not-for-profits (Wardle and Derakhshan, 2017; Graves and Cherubini, 2016). Many work in tandem with Facebook; as of January 2019 some 50 fact-checking groups, who are party to the International Fact Checking Network Code of Principles, independently assess fake news flagged by users (Volpicelli, 2019). The expertise developed includes a variety of flagging and adjudication systems such as NewsGuard’s “nutrition label” that evaluates some 2,000 online news sources, or, as it relates, the sites that garner about 95% of engagement in the news sector (2019).

#### *Automation*

Brief mention should perhaps be made of automation as offering methods for flagging dubious or false content, however much it is rarely recommended in governmental reports. With respect to fact-checking, if there are shared databases of ‘already fact-checked’ stories as well as sources, then software could cross-check suspicious ones against those already debunked or evaluated, as the UK parliamentary report mentions. The discussion concerning the need for human reviewers for content interpretation and curation remains pertinent.

#### *Counter-narratives*

In Germany the government chooses to actively participate in spaces where disinformation is spread. “On these platforms, the German Government provides both reliable information that can be fact-checked and a narrative based on this information” (German Federal Foreign Office, 2018). In that vein, [rumoursaboutgermany.info](http://rumoursaboutgermany.info) is a website for collecting and countering disinformation about Germany spread by human traffickers. While Germany chooses to work with counter-narratives, others have criticised this approach. A Canadian intelligence report argues that developing counter-narratives is a “one event at a time approach” that “fails to address the source and methodology of information campaigns” (Canadian Security Intelligence Service, 2018: 66).

#### *Media literacy and digital ‘hygiene’*

The EU high level expert group on fake news and online disinformation makes a case for increased media and information literacy to counter disinformation, which should be “implemented on a massive scale in school and teacher training curricula” (EU HLG, 2018: 26). This media literacy also should involve the development of tools and training modules for journalists. As a particularly relevant method, the group proposes “more powerful tools to be able to visually map online networks and connections to understand how disinformation is being created, spread and amplified” (EU HLG, 2018: 28).

Some countries speak of 'digital hygiene' when referring to media literacy practices, for instance in France when making a case for the development of skills to assess the validity of the arguments and the reliability of the source. “This is a public hygiene measure—just as people in the 19th century learned to wash their hands” (Jeangène Vilmer et al., 2018: 179). In Sweden the word 'cyberhygien' is employed. The Swedish Civil Contingencies Agency has published a handbook for communicators in public sector organizations for the countering of disinformation, which includes strategies that range from source checking and recognizing a bot to choosing an appropriate response to disinformation. The Swedish Media Council developed a media literacy programme for young people, teaching them critical thinking and disinformation detection; it includes a set of educational materials on ‘source criticism’ ('Källkritik') (Government Offices of Sweden, 2017; Swedish Media Council, 2019). Several recent reports stress the importance of better equipping journalists with tools and skills to recognise and avoid disinformation, mentioning the importance of fact-checking, critical source assessment and ethics (Jeangène Vilmer et al., 2018; Wardle and Derakhshan, 2017).

#### *Investing in civil society and building public trust*

A more general way forward that is presented in the literature is to invest in civil society, as it “must remain the first shield against information manipulation in liberal, democratic societies” (Jeangène Vilmer et al., 2018: 169). Such initiatives are specifically relevant around events such as elections, in which civil society can be supported through non-legislative, pre-emptive measures and multi-stakeholder collaboration of government with the industry, non-governmental sector, and regional actors (Hacıyakupoglu et al., 2018). In Sweden, the aforementioned Swedish Media Council is an example in which politicians and media professionals collaborate and meet regularly to discuss and counter disinformation and related challenges. Such regular, multi-stakeholder consultation both within and across European countries is among the recommendations often given (Brattberg and Mauer, 2018).

#### *Guaranteeing participation in public debate by all*

Lastly is the admonition issued in the 2017 joint UN declaration on “fake news” that emphasised the need for states to enable the participation of all in public debate. They should ensure that any efforts to quell or thwart the practices of fake news-making and spread as well as that of disinformation be handled within the context of the freedom of expression and the freedom on the press (McGonagle, 2017).