

# Fakebook: why Facebook makes the fake news problem inevitable

PAUL BERNAL<sup>1</sup>

*University of East Anglia*

## Abstract

*The current 'fake news' phenomenon is a modern manifestation of something that has existed throughout history. The difference between what happens now and what has happened before is driven by the nature of the internet and social media – and Facebook in particular. Three key strands of Facebook's business model – invading privacy to profile individuals, analysing mass data to profile groups, then algorithmically curating content and targeting individuals and groups for advertising – create a perfect environment for fake news. Proposals to 'deal' with fake news either focus on symptoms or embed us further in the algorithms that create the problem. Whilst we embrace social media, particularly as a route to news, there is little that can be done to reduce the impact of fake news and misinformation. The question is whether the benefits to freedom of expression that social media brings mean that this is a price worth paying.*

**Keywords:** fake news; Facebook; social media; internet; privacy; freedom of expression

## 1 Fake news

On 3 November 2016, Craig Silverman and Lawrence Alexander wrote a piece for BuzzFeed, identifying 'more than 100 pro-Trump websites being run from a single town in the former Yugoslav Republic of Macedonia'.<sup>2</sup> Their headline read 'How teens in the Balkans are duping Trump supporters with fake news'. This appears to have been the first time the words 'fake news' were used – but the phenomenon that it illustrates is something that has existed for a very long time, perhaps as long as history itself.<sup>3</sup> Silverman and Alexander's story covered one very specific form of fake news – websites that purported to be US news and political websites providing information and analysis, but were in fact created in Macedonia and full of stories that, in essence, were not true. The Macedonian teens were doing it entirely for financial gain – they had quickly discovered that they could make advertising money through Google AdSense because of the numbers of

---

1 Dr Paul Bernal, Senior Lecturer in IT, IP and Media Law, UEA Law School.

2 See [www.buzzfeed.com/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo?utm\\_term=.copGOX3Rpo#.jmXg3vJm0R](http://www.buzzfeed.com/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo?utm_term=.copGOX3Rpo#.jmXg3vJm0R).

3 Herodotus, known as the 'father of history' is also known as the 'father of lies' (see e.g. J A S Evans, 'Father of history or father of lies: the reputation of Herodotus' (1968) 64(1) *The Classical Journal* 11–17). Fake news may be even older: there is no real reason to believe that misinformation, disinformation and other phenomena related to 'fake news' did not occur in prehistory, and plenty of reason to believe that they would have.

people visiting their sites. Their ‘fake news’ was popular and successful – and its success provided a number of clues to why the age-old problem of misinformation had gained a new and spectacular lease of life. Those clues need to be understood by those who wish to do something against ‘fake news’ – because they were understood by those who wished to harness its effects and effectiveness. Fake news moved from being a small nuisance, a few Macedonian teenagers making money, to something that might have the potential to aid in the undermining of democracy in many places around the world.

A large part of why this is the case relates to the way that the social media – and Facebook in particular – functions. Though, as shall be shown, fake news has existed for a very long time, with varying degrees of success, the ‘new world’ of social media is ideally suited to its propagation and effectiveness, bringing it to a whole new level. Many different aspects of Facebook in particular contribute to this: the aggregation and analysis of data, the profiling and the targeting. These start with the ‘natural’ tendencies of the internet towards polarisation, where ‘like-minded’ people come together, act as echo chambers and reinforcers – as noted by Sunstein as early as 2001,<sup>4</sup> and taken further by Bernal and Pariser in 2010 and 2011 respectively, through the concepts of ‘back-door Balkanisation’ and ‘filter bubbles’<sup>5</sup> – and magnifies and builds upon them. There are both ‘automatic’ phenomena – a tendency for fake news to spread by its nature and the structures of social media – and deliberately instigated phenomena through which malicious actors can cause fake news to spread for their own purposes, whether these be political, for financial gain (as for the Macedonian teenagers) or simply out of amusement (as for some of the original ‘trickster’ internet trolls).<sup>6</sup>

The way that these phenomena are supported and, in some ways, driven by some of the fundamental aspects of Facebook and other social media means that the current attempts to address the fake news problem are highly unlikely to be successful in any meaningful way. Most address the fake news itself rather than the way that it is spread or the reasons it is effective – and as fake news has always existed and has become remarkably easy to create, particularly in the social media, this is likely to be at best a futile game of whack-a-mole.

If the problems of fake news are really to be addressed, something more fundamental needs to happen. Either the way we use social media has to change – most importantly the use of Facebook and similar social media as a source of news – or the structure and business models of social media have to change fundamentally. The first seems unlikely; the second all-but impossible, as without data analysis, profiling and targeting it is hard to see a business model for Facebook and other social media enterprises. Whilst Facebook and similar social media exist, there will be a problem with fake news – and as the public seems to have little appetite for leaving Facebook, that means that the problems associated with fake news are just going to grow, no matter what lawmakers and technologists try to do about them.

---

4 In Cass Sunstein, *Republic.com* (Princeton University Press 2001) and further developed in Cass Sunstein, *Republic.com 2.0* (Princeton University Press 2007).

5 ‘Back-door Balkanisation’ is discussed in P A Bernal, ‘Web 2.5: the symbiotic web’ (2010) 24 (1) *International Review of Law, Computers and Technology* 25–37, and ‘filter bubbles’ in Eli Pariser, *The Filter Bubble: What The Internet is Hiding from You* (Viking 2011). It should be noted that there has been some questioning of the polarisation effects – see e.g. L Boxell, M Gentzkow and J M Shapiro ‘Is the internet causing political polarization? Evidence from demographics’ (NBER Working Paper Series, Working Paper 23258, 2017) <[www.nber.org/papers/w23258](http://www.nber.org/papers/w23258)>.

6 See P A Bernal, *The Internet, Warts and All: Free Speech, Privacy and Truth* (Cambridge University Press 2018), ch 8, particularly 196–9.

### 1.1 WHAT IS FAKE NEWS?

The concept of ‘fake news’ is somewhat nebulous: a precise definition is neither easy nor necessarily useful. Journalist Matthew D’Ancona defined fake news as ‘the deliberate presentation of falsehood as fact’<sup>7</sup> which captures the ‘fakeness’ but not necessarily the ‘news’ part of the concept. What characterised the stories of the Macedonian teenagers was in part that they were presented as though they were conventional – ‘real’ – news, on sites that looked like news sites. Fake news in its purest form is not just falsehoods pretending to be true (essentially just lies), but falsehoods presenting themselves as both real and ‘news’, in the sense that they are new, relevant and important enough to be ‘newsworthy’.

The nature of how we now consume news means that fake news does not necessarily have to be presented that way. A Facebook page or even a tweet could count – so long as the Facebook page or tweet was both ‘fake’ and ‘news’. Tweeting something untrue about what you had for breakfast would not count, for example, but tweeting that you saw a politician in a compromising situation could: it would be *newsworthy*. Similarly, YouTube videos taking footage of one event and labelling it as another – a protest march in one country at one time presented as coming from another country at another time, protesting about something quite different – would count as fake news.

Concentrating on what exactly constitutes ‘fake news’, however, misses much of the point. The European Commission’s ‘independent High Level Group on fake news and online disinformation’<sup>8</sup> gets part of the way to seeing this. ‘The threat is disinformation, not fake news’ is the headline in its ‘problem definition’ section. It goes on to say that ‘the term is inadequate to capture the complex problem of disinformation, which involves content that is not actually or completely “fake” but fabricated information blended with facts’ and that this includes ‘practices that go well beyond anything resembling “news” to include some forms of automated accounts used for astroturfing, networks of fake followers, fabricated or manipulated videos, targeted advertising, organized trolling, visual memes, and much more’.<sup>9</sup>

Whilst the High Level Group is right that the fake news itself is not the real problem – and that it has always existed – and right that what might be called the disinformation campaigns use a wide range of tactics, it does not look in any depth into the role that the systems and business practices of the social media play in this. The disinformation campaigns are not accidents or exceptions, nor are they the result of rogues misusing systems in a way that they were not intended, but are a ‘natural’ extension of the way that those systems are designed. Data analysis, profiling and targeting in order to persuade and manipulate is precisely the way that Facebook in particular works.

### 1.2 FAKE NARRATIVES

What is also missed in most discussions of fake news – and in the High Level Group’s report – is the role of the conventional media and of what might loosely be described as ‘mainstream’ politicians in disinformation. In particular, in pushing fake *narratives*.<sup>10</sup>

7 See <[www.gq-magazine.co.uk/article/fake-news-inquiry](http://www.gq-magazine.co.uk/article/fake-news-inquiry)> and, for a more general discussion, Matthew D’Ancona, *Post Truth: The New War on Truth and How to Fight Back* (Ebury Press 2017).

8 European Commission HLEG, *A Multi-dimensional Approach to Disinformation* (Report of the Independent High Level Group on Fake News and Online Disinformation 2018) <<https://ec.europa.eu/digital-single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation>>.

9 Ibid 11.

10 See Bernal (n 6) ch 9, 234–9.

Disinformation is not just about individual facts being false, but about the interpretation or analysis of those facts being used to draw false conclusions, to paint false pictures and thus to persuade people to make decisions that have very different effects to the ones they think.

This is not something new: it is a classical tactic of both politics and the media to persuade people to vote against their own interests and in favour of the interests of the politicians or their backers. The misinformation campaigns run on the internet can feed from and at the same time bolster similar campaigns in the conventional media, driven by conventional politicians. In the UK, for example, the idea that immigration is damaging to the economy and to jobs, and a drain on the health service and the welfare system, though contradicted by substantially all the empirical and academic evidence,<sup>11</sup> has been pushed for many years by both the media and politicians. That ‘immigration is damaging’ narrative can then be exaggerated and spread through the online disinformation campaigns – and indeed spread automatically by the social media’s polarisation tendencies.<sup>12</sup>

For the purposes of spreading a fake narrative, it is not just the clearly fake news that can have the effect. ‘Real’ news that fits the same narrative would work just as well, as would information that is neither entirely real nor entirely fake. This mirrors the difference between lying and *bullshitting*, as memorably set out in philosopher Harry Frankfurt’s essay *On Bullshit*.<sup>13</sup> To Frankfurt, lying is saying something you know to be false, while bullshitting is not caring – and perhaps not even knowing – whether what you say is true or false. Those spreading fake narratives are, in Frankfurt’s terms, *bullshitters*, whilst those spreading straightforward fake news are liars. As Frankfurt notes, ‘bullshit is a greater enemy of the truth than lies are’,<sup>14</sup> as those spreading it have no regard for the truth at all. In practical terms it is easier to oppose fake news than fake narratives. Facts can be disproven, perpetrators of lies challenged through laws from fraud to defamation. Narratives, as interpretations of facts, are more like opinions – given much more protection under freedom of speech laws, and quite rightly so. In terms of political manipulation, however, they may be every bit as damaging.

### 1.3 HISTORICAL FAKE NEWS

Though fake news is presented as something ‘new’, there have been forms of it throughout history. Making up stories about current events, or about particular people or groups of people, and using whatever was the most popular or effective method of spreading those stories at the time, appears a ‘natural’ thing to do if you want to persuade people to do something. Using this approach to writing history, for example, has a long history itself. The twelfth-century monk Geoffrey of Monmouth’s ‘The History of the Kings of Britain’ which claimed to trace the lineage of Britons to the Trojan Aeneas, included the mythical King Arthur, creating a narrative of ancient nobility and glory to support his current masters. Evidence to support this narrative was not the point: the effect on the events of the time was.

A closer parallel to the current fake news phenomenon comes from the fifteenth century and the case of Vlad Tepeş: Vlad the Impaler.<sup>15</sup> Vlad, the ruler of the Principality of Wallachia, in current-day Romania, made a large number of enemies through his

11 See <<http://blogs.lse.ac.uk/politicsandpolicy/immigration-demons-and-academic-evidence>>.

12 See Section 2.3 below.

13 Harry G Frankfurt, *On Bullshit* (Princeton University Press 2005).

14 *Ibid* 55.

15 More details of this story can be found in Bernal (n 6) ch 9, 231–4.

various political activities and actions, and those enemies sought to blacken his name at the time, and his memory after he was dead – to damage his image as a hero to many of the people of the principality. The narrative they created was that of Vlad as a vicious, evil and almost demonic figure. The means they used were those current at the time – from manuscripts in Russian Orthodox monasteries and chronicles in the court of the Turkish sultan to records written for the then King of Hungary, Matthias Corvinus. The most interesting and most relevant to the current ‘fake news’ issues was a series of pamphlets prepared by ethnic Germans in the region, between 1488 and 1530. These were particularly gruesome, illustrated with graphic woodcuts. One notable example depicted Vlad eating his dinner with a smile on his face, many people impaled on long sticks in front of him as he ate. Did he actually eat dinner, surrounded by his victims? Was he actually more brutal than other rulers of the times? The facts did not really matter to those presenting them – and when examined by historians from this distance, it appears he was not that remarkable in his brutality but an effective leader and warrior:

Vlad was a man of his times – a capable leader, driven by clear objectives, yet flawed, as men are. Had he been the degenerate, inhuman creature that some have transformed him into, Dracula would never have regained his throne in 1476, supported by the Christian leaders of Southeastern Europe. Unfortunately, the propaganda of the time distorted the image of this brave prince whose role in history deserves reconsideration.<sup>16</sup>

The truth of the details was not important to the German creators and distributors of the pamphlets. They had a narrative to support, a story to spread, and they used the methods of the time. Pamphlets distributed by hand, with graphic images to help those who could not read, were as good and quick a method of spreading a story as was available at the time: the Facebook pages of the fifteenth century. They reached the communities the creators wanted them to reach, in the Saxon towns in Wallachia and Transylvania, provided them with ammunition to push the ideas they supported, the ‘facts’ to ‘prove’ that they were right about Vlad. The pamphlets were part of a bigger phenomenon – one that worked so well that though the facts of his history are known by very few, ‘Vlad the Impaler’ is well ‘known’ as one of the worst villains in history and a byword for brutality.

In a similar way, Richard III had his name thoroughly blackened by Tudor propaganda – the relationship between fake news and propaganda is very close – most directly through Shakespeare’s memorable play. Oliver Cromwell’s reputation was attacked by restoration royalists in amongst other things ‘hostile biographies, written just after the Restoration, which gleefully printed invented, scandalous stories about him in order to blacken his character’.<sup>17</sup> In seventeenth-century France, misinformation was spread using pamphlets called *affiches* and tiny *billets* – card-sized reports that could easily be concealed and then discarded, or sung or spoken by roaming *colporteurs*. As historian Linda Kiernan describes it:

News was relayed (and redrafted) often by word of mouth – the social media of the age. Colporteurs peddled songs and stories on the streets, often scribbled down on notes, or sung aloud to an already well-known tune. Making the

16 Kurt W Treptow, *Vlad III Dracula: The Life and Times of the Historical Dracula* (Center for Romanian Studies 2000) 178, quoted in Bernal (n 6) 234.

17 Barry Coward, *Cromwell: Profiles in Power* (Pearson Education 1991) 7. Amongst other things, stories related to sexual, political, religious and other aspects of Cromwell’s life.

message memorable was half the battle – the other half could be covered by the outlandish or scandalous nature of the material.<sup>18</sup>

The nineteenth century demonstrates that, as in the current era, the motivation for those producing and distributing fake news can be financial as well as the more obviously manipulative. In Germany, a fashion grew for foreign news, driven by technological developments from printing to the telegraph. To differentiate themselves from those using standardised news agency reports, some sent foreign correspondents abroad, to give first-hand reports and provide local details and colour. That cost money, and a new strategy arose: the ‘*unechte Korrespondenz*’ or ‘fake foreign correspondent’s letter’.<sup>19</sup> These used techniques instantly recognisable to current creators and distributors of fake news. McGillen describes how one of the most noted fake foreign correspondents, Theodor Fontane, put together his report on London’s Tooley Street fire in 1861:

He had gleaned lots of passages from older articles about the fire, partially revised them, added a translated section from a current story in the London Times, and had dressed it all up with a feigned first-person perspective and some very creative additions, such as the ‘companion’ who had allegedly helped him pass the police cordon.<sup>20</sup>

The point, as McGillen notes, is to fit in with the expectations of the reader. Use familiar ideas and images, match a narrative that the reader expects and perhaps supports, and add colour and flavour to make it more attractive. It used the most readily available and effective form of communication of the time – just as the enemies of Vlad Tepeş had used the most readily available forms of their time. The motivations for the *unechte Korrespondenz* were at least partially financial, just as those of the Macedonian teens identified by Silverman and Alexander were financial, but they also had a distinctly political edge:

... false correspondences provided a steady trickle of misinformation, with every convincing false account – even those on ‘soft,’ non-political topics – incrementally increasing the readers’ belief in the overall trustworthiness and competence of a given newspaper. This steady work on a paper’s reputation, and the gradual chipping away at the competence and credibility of competing papers, helped prepare the soil on which the less subtle propaganda messages could flourish.<sup>21</sup>

As different media become more important and have bigger audiences and impacts, those wishing to spread disinformation – including specifically fake news – embrace them and adapt to them. William Joyce – Lord Haw-Haw – and the so-called Tokyo Roses – used radio broadcasts in the Second World War, from Germany to the UK and Japan to the USA respectively. Anna Wallis Suh – Seoul City Sue – and Trinh Thi Ngo – Hanoi Hannah – did the same from North Korea and North Vietnam and to US troops in the Korean and Vietnam wars respectively. Muhammad Saeed al-Sahhaf – labelled as Comical Ali – used television interviews during the 2003 war in Iraq. All used combinations of fact and fiction, weaving them together to support a particular narrative and have a political effect. The medium chosen in each case was what was most easily available and most

18 See <<https://theconversation.com/frondeurs-and-fake-news-how-misinformation-ruled-in-17th-century-france-81196>>.

19 See Petra McGillen, ‘From our own false correspondent: armchair reporting in the Kreuzzeitung as a means of early information warfare, ca. 1860–1870’ in Hansjakob Ziemer (ed), *Observing the Everyday: Journalistic Practices and Knowledge Production in the Modern Era* (forthcoming).

20 Ibid.

21 Ibid.

likely to have an effect. In the Second World War and for the troops in the Korean and Vietnam wars, radio worked best. By the time of the invasion of Iraq, it was television. In the current era, it is the social media that is both the most popular and the most likely to have an effect.

## 2 Facebook

Facebook is not only one of the fastest and most efficient forms of communication available, it is also tailor-made for the spreading of fake news and for political manipulation. It brings with it the largest possible audience – well over 2 billion people are now on Facebook. It has also become a key place for accessing news. In 2017, the Pew Research Center reported that 67 per cent of Americans got at least some of their news from social media, whilst 68 per cent of Facebook users used Facebook to access news.<sup>22</sup> This is on the increase, and spreading amongst different and critical demographics: Pew also report a notable increase in the use of social media for news amongst older, non-white and less educated people. In particular, Pew noted that 55 per cent of Americans over 50 reported getting their news from social media sites, an increase of 10 per cent over the previous year. Pew's data concerns Americans, but social media and Facebook in particular are a major source of news through much of the world – including most of the key places for potential political manipulation.<sup>23</sup>

The audience is just the starting point. Just as Facebook combines many different forms of communication – from keeping families updated and friends sharing personal updates to romance, job-searching and information-seeking, from the intimate to the broadcast – it also combines all the key elements for the design, creation, targeting and propagation of fake news.

There are lessons to learn from both the historical examples and the more recent manifestations to show how this works. Firstly, that for fake news to be successful, it has to have a potential audience. Secondly, that it has to match at least some of the expectations of that audience. Thirdly, in terms of measuring what constitutes success, that it has to produce the results that the person creating or spreading the news desires. That may mean producing a political or propaganda success – blackening the name of Vlad the Impaler, Oliver Cromwell, Hillary Clinton or Jeremy Corbyn for example – or generating income, as for the German fake foreign correspondents or the Macedonian teens producing fake stories aiding Trump.

### 2.1 CRAFTING FAKE NEWS ON FACEBOOK

Facebook not only aids those wishing to spread 'fake news', but provides them with tools and incentives to do so. The starting point is the creation of fake news. In the past this was relatively difficult. Neither printing presses nor paper were cheap or easily available until comparatively recently. They were often large and hard to conceal – government censors could find dissident printers and destroy their presses. The stage-by-stage technological developments, from commonly available computers to desktop publishing to webpages to Facebook pages has made this progressively easier – mostly in ways distinctly beneficial to freedom of speech, but also enabling the production of fake news. With Facebook, this has reached a particular extreme: two Facebook pages look much like each other, regardless of whether they are made by a major news provider or a Macedonian teenager. Making your 'fake' news look like the real thing has never been

22 See <[www.journalism.org/2017/09/07/news-use-across-social-media-platforms-2017](http://www.journalism.org/2017/09/07/news-use-across-social-media-platforms-2017)>.

23 See Reuters Institute Digital News Report 2017, Reuters Institute for the Study of Journalism <[www.digitalnewsreport.org](http://www.digitalnewsreport.org)>.

easier. Moreover, the costs involved are negligible – the technology is universal, Facebook is ‘free’ to those using it to create pages, and distribution costs are minimal.

Facebook also makes the crafting of the *content* of fake news easier – and this begins the critical point about its business model. Facebook’s systems for data analysis and for profiling create means for the advertisers (and fake news providers) to discover what people are interested in – both in general terms, following large groups, and in specific terms, profiling individuals. Given the mass of people on Facebook and the mass of data provided by them, the opportunities for analysis are of an unprecedented level and the quality of the profiling is improving all the time. As Kosinski et al revealed through their 2013 study,<sup>24</sup> Facebook ‘likes’ in particular ‘can be used to automatically and accurately predict a range of highly sensitive personal attributes including: sexual orientation, ethnicity, religious and political views, personality traits, intelligence, happiness, use of addictive substances, parental separation, age, and gender’. Specifically, their study was remarkably accurate on race, sexuality and political views, discriminating between ‘homosexual and heterosexual men in 88% of cases, African Americans and Caucasian Americans in 95% of cases, and between Democrat and Republican in 85% of cases’.

This is not through explicit information about these subjects by the relevant people, but through an analysis of their likes, including the most mundane. When the study was released, the headlines included that intelligence could be indicated by ‘liking’ curly fries, and a lack of intelligence by ‘liking’ motorcycles – the enormous amount of data available through Facebook allows seemingly random correlations to be found and used as predictors, through mathematical analysis rather than psychological insight. The sheer scale of Facebook is what gives this its power. The more data that is available about a population, the more easily traits can be determined. The more data that is available about the population *as a whole*, the less data is needed about a particular individual in order to profile and categorise them.

For the malicious crafter of fake news this provides many benefits, but the point at this stage is that it can be used to choose which issues to focus on within the stories. If people are worried about immigration, and about Eastern European immigration in particular, and are worried by particular issues such as the use of health services, then creating a story about this is easy. Through the data-mining and analysis, Facebook can allow people to identify the issues much more precisely than that. Perhaps specifying particular hospitals. Perhaps finding that people are more worried about Bulgarians than Romanians, or Poles rather than Czechs. Trying to make a story that is more likely to be believed and is also more likely to persuade, starts with an analysis of people’s existing beliefs.

## 2.2 DELIVERING FAKE NEWS VIA FACEBOOK

At the same time as enabling the creator of fake news to find what kind of fake news is likely to attract an audience, it immediately helps the fake newsmonger to deliver that fake news to precisely that audience. This would have been far harder even in the recent past: finding people who share particular religious beliefs, for example, or views on politics, was a challenge. Now it is a matter of a few clicks. As noted above, very little mundane information is needed on one particular individual in order to be able to make some specific assessments in very much more intimate areas. In politics and fake news, just as in advertising, it does not matter whether every ‘guess’ is accurate: it is a matter of getting

---

24 Michal Kosinski, David Stillwell and Thore Graepel, ‘Private traits and attributes are predictable from digital records of human behaviour’ (2013) 110(15) *Proceedings of the National Academy of Sciences*.

enough ‘hits’ to produce a result. The rates noted above for assessments of race, sexuality and political leanings are remarkably good in this context.

Facebook does not just help with identifying the potential audience but also provides the mechanisms to target them – from the tools for advertisers to the various groups, pages and so forth where they can be found. Paid political advertisements are only a small part of this – the spreading of information in other ways is a critical part of the way that fake news works. Critically, much of this happens without those behind the fake news needing to take any action: Facebook’s whole ‘sharing’ system does much of this automatically. It is not just that the creators of fake news are able to find an audience, using tools provided by Facebook, but that the consumers of fake news – willing and aware or unwilling and unaware that what they are consuming is fake – are able to find that fake news, also using tools provided by Facebook. ‘News’ tailored to individuals, searching for news on particular subjects (using terms that the providers of fake news will know are likely to be searched for) and so forth all work in the same direction: to ensure that what people are looking for is available to them, and that they can find it. Once people have found the items of fake news, they will share them – and Facebook’s social network will ensure that it is spread without the creators of the fake news having to do any more about it: an electronic grapevine more effective than anything previously imaginable, and one particularly suited to fake news, for a series of reasons.

Perhaps most importantly, fake news can be more believable than ‘real’ news. It fits in with people’s prejudices. This could be by design, for carefully crafted fake news, as noted above. It could be because the person has been seeking out a particular kind of news – searching, for example for ‘Bulgarian immigrant crime’, when they have a preconception that Bulgarian immigrants are committing a lot of crime in the UK. It could also be because Facebook – or Google, Twitter etc. – are pushing those kinds of stories at the people who are predisposed to believe them. Facebook has been tailoring its ‘news feed’ since it was introduced in 2006,<sup>25</sup> and that tailoring has been based on Facebook’s algorithmic assessment of the individual’s preferences. This will automatically *push* stories on the subjects that a person is interested in directly to that person. For the fake news provider, this is perfect: create the fake news, and Facebook will push it to people predisposed to sympathise with the message and indeed to believe the story. In turn, those people will then share it with their own networks of similar-minded people, transfer it from Facebook onto their other social networks and so forth.

With fake news, ‘plot holes’ in stories can be plugged, consistency both internally and with other fake news can be ensured – and as with the work of Fontane, familiar and known ‘facts’ can be put in to help readers along. Well-known places, recognisable public figures and so forth can be strategically placed to help the reader along. Things can be made more dramatic or exciting than ‘real’ news, making people more likely to read and enjoy:

Fontane’s false correspondences at times simply provided the better stories in terms of their dramaturgy, aesthetic and intellectual appeal, and entertainment value.<sup>26</sup>

This makes fake news more likely to be believed *and shared*. There is empirical evidence to support this, from more than one direction, that hits at the heart of the way that Facebook works, and the way that its business model works. What is also clear is that Facebook knows this, and in some ways wants people to know this – because it both commissioned and publicised some of the research that demonstrates it.

25 See the official launch notification of Facebook’s News Feed and Mini Feed <[www.facebook.com/notes/facebook/facebook-gets-a-facelift/2207967130](http://www.facebook.com/notes/facebook/facebook-gets-a-facelift/2207967130)>.

26 McGillen (n 19).

### 2.3 FAKE NEWS AND FACEBOOK'S MODEL

Facebook's controversial 'emotional contagion' experiment from 2014 is perhaps the best known.<sup>27</sup> It involved the manipulation of the news feeds of 689,003 Facebook users in an attempt to manipulate their emotions. Through the use of 'sentiment analysis', they assessed whether items on people's news feeds were happy, sad, or neither happy nor sad. They then divided the group into three and manipulated their news feeds in an attempt to influence their emotions. For one group, they promoted happy posts and suppressed sad ones, for the second they promoted sad posts and suppressed happy ones, and for the third they suppressed both happy and sad posts, leaving the posts without emotional content prominent. Then, using a similar sentiment analysis on subsequent posts by the users, they assessed whether this manipulation of their news feeds had any effect on the users' emotions. The result that seemed to gather the most interest was that, to a statistically significant degree, users whose feeds were made 'happier' became happier, whilst those whose feeds were made 'sadder' became sadder. What was just as interesting, however, was what happened to the third group, given feeds with less emotional content. They engaged less with Facebook. For Facebook, that matters – and it means that Facebook has a direct incentive to prioritise emotional over non-emotional content, and not just in their news feed. It also suggests that users who receive more emotional content, because they are more likely to engage, are also more likely to share that or related content – and is part of what supports the 'automatic' generation of fake news already mentioned and discussed further below. Facebook offered the academics who performed this experiment 'encouragement and support'.<sup>28</sup> It is easy to see how the dissemination of this research is in Facebook's interest: to be able to prove to advertisers that you can have an emotional effect on your users is to demonstrate great power.

The next piece of research, that feeds directly into this, is Blaine and Boyer's 2018 study that showed that people were more likely both to show interest in *and to share and spread* rumours that contain threatening content.<sup>29</sup> Further research shows that simply 'reading a fake news headline once is sufficient to increase perceptions of its accuracy',<sup>30</sup> whilst repeating a story makes it more likely to be believed.<sup>31</sup> Facebook's tailoring of news feeds according to interests plays into both of these effects where interests fit with either a fake narrative or a narrative being played upon by fake news providers. Combine this with the way that communities of like-minded people build on Facebook and other similar social media and you have a ripe and ready environment for the automatic spreading of fake news and for the polarisation and generation of extreme views. These communities also become perfect places for the malicious actors who wish to spread their fake news to 'seed'. Plant an appropriately crafted story in the midst of these communities and they will spread it for you. As before, some of them may be spreading it whilst knowing that it is fake, because they know it will have the kind of political impact

27 Adam D I Kramera, Jamie E Guillory and Jeffrey T Hancock, 'Experimental evidence of massive-scale emotional contagion through social networks' (2014) 111(24) *Proceedings of the National Academy of Sciences*.

28 *Ibid*, in the acknowledgments to the paper.

29 Timothy Blaine and Pascal Boyer, 'Origins of sinister rumors: a preference for threat-related material in the supply and demand of information' (2018) 39(1) *Evolution and Human Behavior* 67–75.

30 G Pennycook, T D Cannon and D G Rand, 'Prior exposure increases perceived accuracy of fake news' (2017) <[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2958246](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2958246)>.

31 See, for example, Lisa K Fazio, Nadia M Brashier, Keith B Payne and Elizabeth J Marsh, 'Knowledge does not protect against illusory truth' (2015) 144(5) *Journal of Experimental Psychology: General* 993–1002. This is part of what is referred to as 'the illusory truth effect'.

they want to see, out of a sense of amusement or for some other reason. Many others, however, will genuinely believe that the fake news is true because the fake news is designed – effectively – to be believable specifically to people like them.

Another key concept comes into play here: astroturfing. Astroturf is fake grass – astroturfing, a term coined by Senator Lloyd Bentsen in relation to the activities of PR firms in the 1990s,<sup>32</sup> is the creation of a fake grassroots movement. This is particularly easy on Facebook – using ‘fake people’ and fake pages to make it look as though there is a big and important movement on the ground, though in reality it is not at all what it seems, but something designed specifically to manipulate.

## 2.4 FINDING FAKE NEWS VIA FACEBOOK

Facebook does not just help those who wish to spread (fake) news, it helps those who seek to *find* news. It provides a number of different routes, some of which have already been noted from the other direction: ‘pushed’ personalised news through the news feed, in a form designed to be as attractive as possible to the user and encouraging the user to use their news feed more and more; the ‘social’ aspects of the network – groups, pages, networks of ‘friends’ – as a route to news; the ‘friends’ themselves, whose ‘Mini Feeds’ (how friends see their timelines) are also curated by Facebook in a way to make those friends seem as attractive as possible to like-minded people;<sup>33</sup> and direct searching for news, yet again algorithmically curated by Facebook.

All of this is designed to help users find the kind of stories they want to read – and there is empirical evidence to suggest that this can lead directly to the spread of fake news. This is Kahan: ‘[a] considerable body of research concludes that people’s cultural and political predispositions are the source, not the outcome, of the information they consume. Identity protection, not correction, is their goal . . . armed with evidence, people are less vulnerable to succumb to opposing arguments’.<sup>34</sup> On social media people want to bolster their positions, not to find the truth – and to ‘prove’ that they are in the right. Both Kahan and Pennycook and Rand also suggest that more analytical or numerate people are in some ways both particularly vulnerable to fake news and particularly likely to find it (using the various tools provided by Facebook and others) and hence spread it. This is Pennycook and Rand:

. . . the disposition to think analytically leads people to use deliberative reasoning to justify their prior beliefs and protect their political (or otherwise) identity.<sup>35</sup>

This is Kahan:

. . . the more numerate people are even more likely than the least numerate ones to construe such evidence as supporting the factual beliefs that prevail among people who share their political identity no matter what its true import.<sup>36</sup>

That is, numerate and analytically minded people will have both the propensity to believe news supporting their beliefs (whether it is ‘real’ or ‘fake’ news) and have the ability to

32 See e.g. Thomas P Lyon and John W Maxwell, ‘Astroturf, interest group lobbying and corporate strategy’ (2004) 13(4) *Journal of Economics and Management Strategy* 561–97.

33 Mini Feed, like ‘News Feed’, has been tailored since its launch in 2006. See n 25.

34 Dan M Kahan, ‘Misinformation and identity-protective cognition’ (2 October 2017) 4 <<https://ssrn.com/abstract=3046603>>.

35 G Pennycook and D G Rand, ‘Who falls for fake news? The roles of analytic thinking, motivated reasoning, political ideology, and bullshit receptivity’ (Yale University Working Paper, September 2017) <[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3023545](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3023545)>.

36 Kahan (n 34) 2, citing Dan M Kahan, E Peters, E C Dawson and P Slovic, ‘Motivated numeracy and enlightened self-government’ (2017) 1 *Behavioural Public Policy* 54–86.

find it and use it in a logical way to ‘win’ arguments. ‘Winning’ arguments is one of the keys to political interactions on social media – and Facebook and Twitter in particular. Moreover, there is a strong community of numerate and analytically minded people on the internet, and in particular amongst those very active on social media: the role of the gaming and hacking communities and others with a strong interest in computing is one of critical importance. The popularity of conspiracy theories, for example, in many areas of the internet, fits well with this observation. Rational arguments can be constructed for some of the most bizarre of ideas – and ‘evidence’ found easily and quickly to support it. That evidence then spreads, as do the ‘news’ stories themselves.

Facebook aids and abets the spread of fake news through its basic principles – data-mining and analysis, profiling, personalising, targeting and the encouragement of sharing are all direct supporters of the spread of fake news and the bolstering of fake narratives. It is through Facebook’s basic systems that fake news has become a qualitatively different and quantitatively more significant problem – not just through the actions of rogue actors who can be hunted down and removed. Those who spread fake news on Facebook are the ordinary users, and for a wide variety of reasons. Many, and perhaps most, of those reasons are nothing to do with conspiracies or an attempt to undermine the truth or the political system: people are simply sharing stories that interest them or that they think are important. They may share them maliciously – but they may also share them just because they want to keep their ‘friends’ and other connections informed. They may want to help people to see what they consider the truth – or to help them in their arguments with their ‘opponents’ – conservatives arming other conservatives for their battles with liberals; and liberals arming other liberals for their battles with conservatives. It is part of the polarisation process – and the social media, and Facebook in particular, is perfectly designed to help it to happen.<sup>37</sup>

## 2.5 FACEBOOK AND POLITICS

That social media is well suited for ‘real world’ politics might have come as a bit of a surprise to some of the early adopters of the internet – John Perry Barlow’s 1996 Declaration of Independence of Cyberspace<sup>38</sup> asserted that ‘old world’ politics had neither the place nor the power to have any influence over the internet. Many believed him then and some still believe him now – but the influence of politics over the internet and the ability for ‘real world’ politicians and political movements to use the internet for their own purposes has become all too clear in recent years. This is particularly true for the social media – and both the politicians and the social media companies have seen this as a great opportunity and seized it with both hands.

Social media gives politicians a chance to reach and influence their supporters and potential voters – and increasingly as an analytical tool to understand and target those potential voters. The social media companies see it as an opportunity to make money. As Kreiss and McGregor note in their detailed analysis of how the technology firms worked during the 2016 US presidential election:

. . . the growth of their work in electoral politics was driven by the desire for direct revenues from their services and products, for candidates to give their

<sup>37</sup> See n 5 above.

<sup>38</sup> John Perry Barlow, ‘A declaration of the independence of cyberspace’ (Electronic Frontier Foundation 1996) <[www.eff.org/cyberspace-independence](http://www.eff.org/cyberspace-independence)>.

services and platforms greater public visibility, and to establish relationships with legislators.<sup>39</sup>

It is at least on the surface a mutually beneficial relationship – but one with consequences. Politicians use the analytical tools and immense amount of data available in order to shape their campaigns – paying the technology companies for the opportunity. The technology companies go much further, in working with the political campaigns directly. They have developed specifically partisan teams to do so. This is Kreiss and McGregor again:

Microsoft, Google, Twitter, and Facebook all came to adopt and currently have organizational structures and staffing patterns that are organized along partisan lines. These firms have partisan teams, often made up of practitioners with backgrounds in Democratic and Republican politics, which work with campaigns and parties of the same political affiliation.

This is more than just letting politicians use the services and systems: it is actively encouraging them to do so, developing products and services specifically to help them. On the surface this does not necessarily seem to be a problem – it could look as a positive contribution to democratic processes. As recent events in politics – notably the Brexit referendum and the election of Donald Trump – have shown, the relationship between politics and truth is often troublesome. The use of Facebook and other social media, with their polarising effects and possibilities for fake news and other misinformation, brings this particularly into focus. Having partisan teams driving further uses of social media for politics brings these negative possibilities out even further.

What is abundantly clear is that Facebook, Twitter and other forms of social media are now part of the political sphere. Politics *does* happen on the social media – and this is not a genie that can be easily coaxed back into its bottle. That means that all those who want to be involved in political activities will look at the possibilities – including those who wish to use fake news for these purposes. Boundaries can be and are blurred between what counts as political and what does not. When does an individual sharing an opinion or spreading a new story (real or fake) become a political act? Campaigning groups of various kinds all use the social media – and Facebook in particular – in similar ways to direct and official political campaigns. Tools developed for the ‘real’ politicians can easily be used by the various shades of grey between them and the worst of the political manipulators. There is a financial incentive to the social media companies not only to develop those tools but to develop the markets for those tools. This is just one of the ways that they make money from fake news and related forms of misinformation: it is an incentive that is not going to easily disappear.

## 2.6 FACEBOOK, FAKE NEWS AND THE TRADITIONAL MEDIA

At one level, the traditional media seeks to portray itself as providing the ‘real’ news, as opposed to the fake stuff to be found on the internet. At another level, it has played its part in misinformation and disinformation campaigns throughout the ages – Fontane worked for the mainstream media in Germany, for example, whilst in the UK the print media is notably and unashamedly partisan and has pushed fake narratives and attempted to influence politics on a significant scale. Whilst the broadcast media in the UK is required by law to present its news with ‘due impartiality’ and ‘due accuracy’,<sup>40</sup> the extent

39 Daniel Kreiss and Shannon McGregor, ‘Technology firms shape political communication: the work of Microsoft, Facebook, Twitter, and Google with campaigns during the 2016 US presidential cycle’ (2018) 35(2) *Political Communication* 155–77 doi: 10.1080/10584609.2017.1364814.

40 The Ofcom Broadcasting Code, s 5 <[www.ofcom.org.uk/tv-radio-and-on-demand/broadcast-codes/broadcast-code/section-five-due-impartiality-accuracy](http://www.ofcom.org.uk/tv-radio-and-on-demand/broadcast-codes/broadcast-code/section-five-due-impartiality-accuracy)>.

to which it achieves that is often open to question. In a content analysis in 2013, the BBC's coverage of politics was shown to be dominated by the ruling party's voices.<sup>41</sup> Certain particular topics were framed in ways that could be seen as distinctly partial: coverage of the UK's relationship to Europe was 'frequently framed as a problem' whilst the debate over immigration was 'framed by politicians, whose statements were often presented as "facts"'.<sup>42</sup> In both of these areas, this plays directly into the arguably fake narratives pushed by the partisan print media.

The relationship between the traditional media and the social media is similarly not simple: the two are not separate and offering alternatives, but intrinsically and inextricably interlinked. The traditional media uses the social media as a way to reach its audience, with pages on Facebook not just for the media operations themselves but for the individual programmes and journalists. It is a rare journalist that does not have a presence on Facebook and an active Twitter account. Links to key stories are posted on Facebook and Twitter, images of front pages sent around and much more. Traffic is driven to the websites of the relevant news sources, potentially providing advertising revenue for the media. The social media similarly benefits through traffic and advertising revenues, as well as through the data gathered and profiling made possible through the users' access to news sources. Potentially even more importantly, Facebook gains by becoming even more *the* place to go on the internet: not just for sharing family photographs and a few personal stories, but for news, for politics and more. Facebook becomes part of the infrastructure for everyone.

For the news media it is a feedback loop. If Facebook is the most important way people consume news media, then the media must be on Facebook, which makes Facebook the best place to consume news media. Though that might seem to be mutually beneficial, there are a number of potentially negative consequences for the news media. It creates a dependency, which puts the news media in a position of weakness in relation to Facebook. Facebook can change how it deals with the media – terms and conditions, how it curates news and so forth – and there is very little that the news media can do about it. It also means that the more that news media is consumed through Facebook, the more Facebook takes the advertising revenue from news consumption. Thirdly, news *stories* can be spread individually and selectively, meaning that the news media editorial teams have less control, and news can be seen *and shared* out of context, making it easier for others to manipulate how news is consumed, to use it to shape their own narratives.

The narratives may be the most important point. A fake narrative may be more damaging than fake news – and this is another key angle to the relationship between the traditional and social media worlds. Narratives created by the traditional media can be fed and fostered in the social media – and form the basis for fake news that uses both the stories in the traditional media and purely invented material, further feeding the narratives. In its turn, the traditional media feeds off the social media, using it as sources and attempting to harness its power. It is a kind of symbiotic relationship, though which is the more parasitic of the partners is sometimes hard to tell.

For specific pieces of fake news there are legal and regulatory mechanisms in place for the traditional media – in the UK, including the self-regulation of the press via the Independent Press Standards Organisation and IMPRESS and the statutory regulation of

41 Karin Wahl-Jorgensen, Richard Sambrook, Mike Berry, Kerry Moore, Lucy Bennett, Jonathan Cable, Inaki Garcia-Blanco, Jenny Kidd, Lina Dencik and Arne Hintz, 'Content Analysis: BBC Breadth of Opinion Review' (Cardiff School of Journalism, Media and Cultural Studies 2013) <[http://downloads.bbc.co.uk/bbctrust/assets/files/pdf/our\\_work/breadth\\_opinion/content\\_analysis.pdf](http://downloads.bbc.co.uk/bbctrust/assets/files/pdf/our_work/breadth_opinion/content_analysis.pdf)>.

42 Ibid 5.

the broadcast media through Ofcom and the Communications Act 2003. Errors can be ‘corrected’ – though the form of the correction is often far less prominent than the original story. Corrections to false stories in the media are even less likely to have an effect through Facebook and the rest of the social media: the corrections will not be spread as the original stories were and will not be found by those looking for confirmation of their prejudices and as support for their arguments, rendering this kind of way of dealing with problematic – and effectively fake – news almost entirely ineffective for those who access their news through social media.

### 3 Solutions and conclusions

Many of the legal and regulatory solutions offered to the fake news problems in the new environment are direct descendants of ideas that existed in the old environment: fact-checking; complaints mechanisms; obligations on ‘platforms’ to check and remove fake news. Laws against fake news have existed in many places for a very long time indeed: in England, the First Statute of Westminster in 1275 included ‘none be so hardy to tell or publish any false news or tales, whereby discord, or occasion of discord or slander may grow between the King and his people, or the great men of the realm’.<sup>43</sup> That statute laid the responsibility on the person who ‘was the first author of the tale’ – and to an extent the same pressure is still being suggested today. Creators of fake news are to be rooted out – and the responsibility, both legal and in how much of the debate is being framed, placed upon them. ‘Russian trolls’ or Macedonian teenagers are suggested as the basis of the problem – and there are creators of fake news who do fit that stereotype. Pressure is put on the social media companies both to ban these kinds of creators and to take down the news that they create.

#### 3.1 PROBLEMATIC SOLUTIONS

There are a number of immediate problems with this kind of solution. The first is locating these creators – though not insoluble, it is not necessarily easy. The second is that there is no shortage of people like this. Shut down one and another will immediately take their place. Those who benefit from the results of the disinformation will simply get others to do the job. It is easy to do, quick and relatively cheap at all stages – from the creation of identities to the building of website and Facebook pages. In the past, printing a pamphlet required a bulky printing press, and distributing that pamphlet required a network of trusted individuals. Presses could be found and smashed, networks infiltrated and key members arrested or worse. Now there are few such problems: the mass of the population of most nations have sufficient technology (a PC or a smartphone) to produce fake news, whilst distribution is largely automatic. The people who share fake news are not part of a secret conspiracy, they are just using social media as it is intended to be used: sharing links and stories, engaging in political debate and so forth.

Addressing fake news at its source is therefore, though useful in some ways, unlikely to do more than scratch at the surface of the problem – and is likely to have unforeseen consequences that may cause more problems than they solve. Governments can – and already seem to be – use dealing with fake news as a way to introduce censorship and to silence dissent or opposition.<sup>44</sup> Enemies of a particular perspective will use the weapons

43 See for example <<https://babel.hathitrust.org/cgi/pt?id=pst.000017915496;view=1up;seq=157>> 35.

44 See e.g. <[www.theguardian.com/media/2018/apr/24/global-crackdown-on-fake-news-raises-censorship-concerns](http://www.theguardian.com/media/2018/apr/24/global-crackdown-on-fake-news-raises-censorship-concerns)>, mentioning approaches in France and Germany, and more directly Malaysia and Thailand. The Malaysian law, the Anti-Fake News Act 2018, has been particularly criticised by human rights groups. See e.g. <[www.amnesty.org/en/latest/news/2018/04/malaysia-fake-news-bill-hastily-approved-amid-outcry](http://www.amnesty.org/en/latest/news/2018/04/malaysia-fake-news-bill-hastily-approved-amid-outcry)>.

created to fight such fake news to shut down or undermine their opponents: the way that Donald Trump labels the people in the mainstream media who oppose him as ‘fake news’ is just one example.<sup>45</sup> In trolling ‘flame wars’ it is a common tactic: trolls report their victims as trolls as a way to getting them banned or entangled in regulatory messes. The same will be true of any mechanisms put forward to tackle fake news in this way.

Labelling potential fake news as unreliable is unlikely to help either. Highlighting potentially untrustworthy headlines may make them more likely to be read, and as noted above<sup>46</sup> even reading a headline makes people more likely to believe it. Facebook abandoned its ‘fake news labelling’ system in December 2017, noting:

Academic research on correcting misinformation has shown that putting a strong image, like a red flag, next to an article may actually entrench deeply held beliefs – the opposite effect to what we intended.<sup>47</sup>

Instead Facebook put in place a ‘related stories’ system, trying to put particular news into context – but again it is hard to see how this will have any significant impact. Facebook cites its own research that doing so produces ‘fewer shares’, but that does not say anything except that it is less damaging than the labelling system. Moreover, it gives another route for fake news creators to approach: spread variants of the same story and they can all be put in place by Facebook’s ‘related stories’ system, making it look even more as though there is a big underlying story supporting their narrative.

Other proposed solutions to the fake news problems will suffer similar fates. The idea that Facebook or Google should use artificial intelligence (AI) to detect fake news, then use their algorithms to demote it in their search results or news feeds, making it less prominent,<sup>48</sup> is attractive on the surface, but is also flawed. As Anjana Susarla notes, ‘[d]espite some basic potential flaws, AI can be a useful tool for spotting online propaganda – but it can also be startlingly good at creating misleading material’.<sup>49</sup> Susarla goes on to say, ‘using AI to detect fake news . . . puts technology in an arms race with itself’. The creators and spreaders of fake news learn the tactics and methods of those trying to block them, then turn those tactics on their heads, designing news that will not only beat the ‘fake news detectors’ but get their own news to the top of the pile. Algorithms can be and are gamed in this way in other contexts – the search engine optimisation industry works in precisely this way – and the kind of people wishing to spread fake news tend to have exactly the kinds of skills to do this, or the resources to employ those who do. In a technological arms race, ‘truth’ is highly unlikely to be a winner. When Facebook, facing criticism from conservatives that its ‘trending news team’ of human editors was biased against them, fired the human editors in 2016 and let the algorithms do the work: the result was predictably bad, with fake news rising right to the top.<sup>50</sup> Two years later, Facebook abandoned its trending news project entirely.<sup>51</sup>

All this, however, is missing the key point. It is the very nature of the way that social media, and Facebook in particular, enables fake news and related misinformation to have its

45 Echoing some of the tactics of Fontane, who sought to denigrate his ‘real’ competitors. See McGillen (n 19).

46 See n 30.

47 See <<https://newsroom.fb.com/news/2017/12/news-feed-fyi-updates-in-our-fight-against-misinformation>>.

48 As suggested, for example, by Mark Zuckerberg in his hearing before a joint congressional committee in April 2018. See e.g. <[www.washingtonpost.com/news/the-switch/wp/2018/04/11/ai-will-solve-facebooks-most-vexing-problems-mark-zuckerberg-says-just-dont-ask-when-or-how](http://www.washingtonpost.com/news/the-switch/wp/2018/04/11/ai-will-solve-facebooks-most-vexing-problems-mark-zuckerberg-says-just-dont-ask-when-or-how)>.

49 See <<https://theconversation.com/how-artificial-intelligence-can-detect-and-create-fake-news-95404>>.

50 See e.g. <[www.theguardian.com/technology/2016/aug/29/facebook-fires-trending-topics-team-algorithm](http://www.theguardian.com/technology/2016/aug/29/facebook-fires-trending-topics-team-algorithm)>.

51 See <<https://newsroom.fb.com/news/2018/06/removing-trending>>.

effect. If specific fake news is stopped – which in itself is highly unlikely – it will not stop the spreading of fake narratives, the construction of further fake news, and the use of individual ‘true’ stories out of context to manipulate and obscure. Truth, insofar as it is possible to identify, becomes lost in the spreading morass of stories, shaped and crafted both by malicious individuals and by unknowing groups coming together to support their beliefs.

What methods could actually work are another matter. Whilst Facebook and similar mechanisms are used as a primary source of news it is hard to see how – and it is important to recognise that though Facebook is the most prominent, it is far from the only network or system involved. Twitter is used to spread stories, Instagram to illustrate them, and YouTube to provide ‘evidence’ – the way that ‘false flag’ and similar stories were spread through YouTube and via Google immediately after the 2017 shootings in Las Vegas and Texas are cases in point.<sup>52</sup> YouTube is a key location for conspiracy theorists: there is a large community of ‘flat-earthers’, for example.<sup>53</sup> They all have the same symptoms: ease of creation, algorithmic curation, community ‘sharing’ and targeting. Again, it is not just that fake news will always find an audience, but that the audience will always find fake news.

### 3.2 FAKEBOOK

None of this should be very surprising. Truth was never a key point for Facebook. It was born from the idea of presenting yourself in positive ways – the college ‘facebook’ was always much more about making an impression than in revealing an underlying truth. Its business essence is about advertising – and truth has never been the key to advertising either. What is more, Facebook’s advertising model is based on data-mining, profiling and targeting, three things that underpin the effective creation and distribution of fake news, both for political manipulation and for financial gain. Facebook’s attraction to users is based on sharing, and on finding people with common interests – two more keys to the effectiveness of fake news, fake narratives and related forms of misinformation. Unless Facebook breaks both its business model for advertisers and its attraction for users, therefore, it will remain not just ripe territory and a magnet for fake news, but will actively aid in both the generation and spread of this news.

Facebook has no particular interest in truth. If spreading true stories makes Facebook money, or being known for truthfulness makes it money, then it will do both. If spreading fake news or other forms of misinformation (or allowing that fake news or misinformation to be spread) makes Facebook more money, then it is its interest to do that. More importantly, if measures that might actually have an impact on the spreading of fake news, or other political manipulation, would mean Facebook made less money, then it would be in its interest *not* to take those measures. The evidence so far suggests both of these are true. That is, allowing fake news and political misinformation makes it more money than focusing on the truth.

Whether anything can realistically be done to reduce the political impact of fake news and related misinformation in the social media era is another matter. There are some possible approaches that might have some impact. Reducing the amount that people get news from Facebook and similar sources might be a key. There are some signs this might be happening: there are indications that people are looking for different sources. Reuters

---

52 See e.g. <[www.theguardian.com/us-news/2017/oct/02/las-vegas-shooting-facebook-google-fake-news-shooter](http://www.theguardian.com/us-news/2017/oct/02/las-vegas-shooting-facebook-google-fake-news-shooter)> and <[www.theguardian.com/us-news/2017/nov/06/google-youtube-texas-shooting-fake-news](http://www.theguardian.com/us-news/2017/nov/06/google-youtube-texas-shooting-fake-news)>.

53 See e.g. [www.newyorker.com/science/elements/looking-for-life-on-a-flat-earth](http://www.newyorker.com/science/elements/looking-for-life-on-a-flat-earth)>.

reports a rise in the use of messaging applications for news<sup>54</sup> – and where those applications use encryption this could reduce both the level of surveillance and of gaming of news. It will not, however, reduce the network and polarisation effects, the automatic spreading of fake news and so forth. Further, as Reuters also notes, Facebook ‘dominates both social networking and messaging’ – owning both Whatsapp and Messenger, the two most popular messaging apps.

That might provide part of the answer. Reducing dependency on a particular source, supporting rather than undermining encryption, limiting data-sharing and data-mining to make profiling and targeting harder, as well as a significant effort into public awareness of the role that the social media plays in the spread of fake news and other misinformation could be part of a way forward. All that does not seem likely to make a significant difference so long as we as a culture continue our embrace of systems like Facebook. Whilst we do so, fake news will remain an inevitable problem – and a problem that is likely to increase. The question is whether the benefits of our embracing social media – to freedom of expression in particular – make it a problem that is worth living with.

At every age, technological developments have brought advances in freedom of speech – and at the same time allowed or supported damaging uses of that freedom of speech. It was true for the pamphleteers, the newspapers, for radio and television – and it is true for Facebook and other social media now. Whilst there is money to be made – both by the fake news providers and by Facebook – and whilst there’s political gain to be made by those who desire it, there is no reason to believe that the problems caused by fake news will lessen, and many reasons to believe that they will continue. Facebook’s business model ensures that the first will remain true, whilst human and political nature mean that the latter will never go away. It is not possible to find a ‘clean’ solution. Instead, messy, imperfect ways ahead may be the best way forward. This is not a problem that is going to go away any time soon.

---

54 Reuters (n 23) 13.