Weeding out Fake News
An Approach to Social Media Regulation

Konrad Niklewicz
Weeding out Fake News
An Approach to Social Media Regulation

Konrad Niklewicz
# Table of Contents

- About the Martens Centre .......................................................... 04
- About the author ........................................................................ 06
- Executive summary .................................................................... 08
- The importance of social media .................................................. 10
- Social media are good, but… ....................................................... 14
  - Fake news ................................................................................ 15
    - Definition of the phenomenon .............................................. 15
    - The history of fake news ...................................................... 16
  - Trapped in bubbles .................................................................. 17
  - Bots: social media machine guns ............................................. 18
- How fake news and online bubbles impact politics ......................... 20
- The current regulatory framework for social media ......................... 28
- The industry’s voluntary measures to tame fake news .................... 32
  - Flagging, reporting and debunking ......................................... 33
  - The limits of self-imposed measures ......................................... 36
- A new approach to the problem of fake news ................................. 38
  - Top-down regulation imposed so far ....................................... 40
  - A new proposal for the short and medium term: the ‘notice and correct’ procedure ........................................... 41
- What are the risks attached to the use of press laws? ....................... 43
- Where to regulate? ....................................................................... 46
- A long-term proposal: education .................................................. 48
- Conclusions and recommendations ............................................... 50
- Bibliography ................................................................................ 54

**Keywords**  Social media – Regulation – Communication – Free speech – Press laws
About the Martens Centre
The Wilfried Martens Centre for European Studies, established in 2007, is the political foundation and think tank of the European People’s Party (EPP). The Martens Centre embodies a pan-European mindset, promoting Christian Democrat, conservative and like-minded political values. It serves as a framework for national political foundations linked to member parties of the EPP. It currently has 31 member foundations and three permanent guest foundations in 24 EU and non-EU countries. The Martens Centre takes part in the preparation of EPP programmes and policy documents. It organises seminars and training on EU policies and on the process of European integration.

The Martens Centre also contributes to formulating EU and national public policies. It produces research studies and books, electronic newsletters, policy briefs and the twice-yearly *European View* journal. Its research activities are divided into six clusters: party structures and EU institutions, economic and social policies, EU foreign policy, environment and energy, values and religion, and new societal challenges. Through its papers, conferences, authors’ dinners and website, the Martens Centre offers a platform for discussion among experts, politicians, policymakers and the European public.
About the author
Konrad Niklewicz is a Ph.D. graduate in the field of the social sciences (University of Warsaw, Faculty of Political Sciences and Journalism). He is currently the managing deputy director of the Civic Institute, a Civic Platform think tank. In 2017 he was a Martens Centre visiting fellow. He has previously served as spokesperson for the Polish presidency of the Council of the European Union and as undersecretary of state at the Ministry of Regional Development. Before joining the government, he was both a journalist and an editor for the foreign/economic section of Gazeta Wyborcza, the leading Polish daily newspaper. Between 2005 and 2007 he was the journal's Brussels correspondent.
Executive summary
Social media are becoming the dominant source of information for significant parts of our societies. There are numerous positive aspects of these media, such as their ability to mobilise for a political cause. No one can deny that social media strengthen free speech in general, allow greater and quicker flows of ideas across societies, and add to the quality of life. Yet at the same time, social media may sometimes negatively impact the public debate. This paper analyses how social media platforms influence democracy in Western countries and makes recommendations on how to address the risks that arise effectively.

The first section describes the phenomena of fake news, echo chambers and social bots. The paper then discusses the ways in which fake news, bots and bubbles impact the public debate. Examples from the recent past (the US, French and German elections and the UK referendum) are debated. In the third part, the paper outlines the current regulatory framework in which social media operate, both in the EU member states and in the US. The peculiar status of ‘Internet intermediaries’ is also analysed.

The following section discusses the different voluntary measures self-imposed by social media companies to eradicate fake news from their feeds. Their effectiveness is assessed.

In the fifth section, a novel way of fighting fake news is introduced. The author suggests that social media platforms should be considered media companies and that they should be regulated by modified versions of existing press laws, adapted to suit the new technology. The creation of a ‘notice and correct’ procedure, as it is tentatively called, would provide an effective tool to stop lies from spreading, allowing affected parties, public or private, to protect their rights. By making the social media platforms jointly responsible for the content they publish, governments would create the right incentives for companies to adapt their business models and to modify the construction of their algorithms and policies. The concept of a ‘notice and correct’ procedure is discussed in the context of the freedom of speech: the risks and challenges are analysed. It is underlined that any attempt at censorship must not be tolerated.

In the final section, the paper discusses the improvement of e-literacy as an additional, viable and long-term solution to the problem of fake news. It concludes by identifying the right conditions under which the ‘notice and correct’ procedure could be implemented.
The importance of social media
The digital revolution has reached a critical point: it is clear that digital issues are profoundly changing our societies. In no other environment is this revolution more visible than in communication: no medium in human history has been as powerful as Facebook, Twitter, Snapchat or YouTube are today. These online platforms, which allow Internet users to publish and share emotions, opinions and information, have seen global take-up on an unprecedented scale. In 2012 these platforms combined had approximately 1.47 billion users worldwide. It is estimated that by the end of 2017 they will clock 2.55 billion users, a third of the earth’s population.

Long considered as useful for entertainment only, social media are becoming the dominant source of news and political tools. This is best seen among the younger generation. In 2016 roughly 50% of Americans aged 18–29 used online platforms as their primary source of news. Only 27% watched the news on television, and as few as 5% read print newspapers. Similar trends can be seen in Europe. In a study conducted in autumn 2015, half of all Europeans declared that they used social media at least once a week, 15% more than in 2011. More than 63% of Germans aged 16–18 find out the bulk of the news from social media platforms. According to a journalist interviewed for this paper, the average person checks the news on his or her smartphone via social media applications. People quickly scroll down the so-called newsfeed, ignoring the actual sources of the information. In certain segments of society, social media wield an effective monopoly of information. ‘I wouldn’t know a lot of the news if I didn’t go on Facebook and just look through it’—explained a 16-year-old American girl, interviewed for a study by the US-based Data & Society Research Institute.

---

1 M. Boni, MEP, former Polish Minister of Administration and Digitisation, interview with the author, Brussels, 28 April 2017.
2 As well as Facebook, Snapchat and YouTube, there are many other social media platforms worth mentioning, including LinkedIn, Google+, Pinterest, WhatsApp, Instagram, Tumblr, Flickr, Reddit and Quora.
6 European Commission, Standard Eurobarometer 84, Media Use in the European Union, Autumn 2015, 16.
7 Bitkom, Bitkom Studie: Jung und vernetzt (Berlin, 2014), 19.
8 P. Laloux, interview with the author, Brussels, 13 March 2017.
There are many explanations for the growing popularity of social media. Platforms share one distinctive feature: unlike the press or television, they allow people to indulge in ‘word-of-mouth’, or more to the point, ‘word-of-click’. By allowing people to interact and share views, emotions and information, Facebook, Snapchat and Twitter kill two birds with one stone: they are both attractive and influential. Interpersonal communication (‘gossip’ in plain English) is more than capable of reinforcing or contradicting the messages flowing from other sources.\textsuperscript{11} Moreover, citizens engage in interpersonal conversations at least as often as they watch TV news or read newspapers.\textsuperscript{12}

Social media are not only changing the global media landscape. They are also gaining importance in political marketing. Facebook, for example, can be used to understand voters’ social behaviour better by analysing users’ recorded online activity. In turn, this can lead to an understanding of people’s political behaviour. Furthermore, for a fee, Facebook allows tailored content to be distributed directly to a chosen audience. ‘Online is not at the kids’ table anymore. It is fully integrated in the campaign’, noted Daphne Wolter, the media policy coordinator at the Konrad Adenauer Stiftung.\textsuperscript{13}

The Cambridge Analytica consultancy, hired by both the Leave campaign in Britain during the EU referendum and Donald Trump’s electoral platform, is reported to have used Facebook to build detailed psychological profiles representing 230 million adult Americans.\textsuperscript{14} Once the target groups—or even individuals—have been identified, social media platforms are able to send them content matching their preferences and emotions in an unlimited number of alternative versions. According to press reports, on one particular day of the US election campaign, the Trump team published over 100,000 variations of an online ad that discouraged black and/or Hispanic voters from voting.\textsuperscript{15}

Social media companies claim to have no political, social or any other public agenda of their own. Facebook, the biggest of them all in terms of revenue and the number of users, portrays itself as a neutral platform that allows people to connect with each other, to share feelings, to inform and be in-

\textsuperscript{12} Ibid., 3180.
\textsuperscript{13} D. Wolter, interview with the author, Berlin, 7 March 2017.
\textsuperscript{15} Ibid.
formed, and to entertain and be entertained. Yet time and again it has been claimed that Facebook has interfered with content published on the platform.\textsuperscript{16} One thing seems certain: over the years, its profile, like those of its competitors, has changed. Today social media platforms more closely resemble media companies. As is explained later, this entails tremendous consequences.

Social media are good, but...
Social media reinforce free speech and greatly induce and facilitate the flow of ideas across different (geographical, societal etc.) borders. They help us in our daily lives and they are an important part of what constitutes our civilisation today. In numerous cases social media have played a positive role in political and social developments, allowing, for example, for the greater mobilisation of pro-democratic movements (as in Egypt in 2011 or Ukraine in 2013–14). The business models behind social media platforms offer unique examples of ground-breaking commercial and technological innovation combined with visionary entrepreneurship. They are part of the great technological change we are witnessing across the globe. As such, they merit praise.

But there are at least three serious problems related to Facebook, Twitter, YouTube and the other online giants, as discussed below. These problems directly impact the way in which Western democracies work.

Fake news

Definition of the phenomenon

The sharing of fake news on social media is deliberate, undertaken by an organisation or an individual, with the aim of fabricating and disseminating information that is fully or partially false in nature in order to influence opinion or stir controversy, or for financial gain. Fake news often includes a grain of truth, but this ‘kernel of true information’ is twisted, taken out of context, surrounded by false details and so on. In many cases, fake news is dressed up to look like a genuine news item, sometimes from a pretend news outlet. Some fake news sources imitate trustworthy, independent institutions. Not all fake news authors aim to make people change their minds. Sometimes the goal is to polarise society.

Fake news must be distinguished from satirical content: making up absurd details and claims does not meet the definition of satire. Online fake news gains traction when it is shared massively on Facebook, Twitter, Snapchat and/or WhatsApp, or is promoted by Google’s search algorithms. This is how such items infiltrate the political debate. One of the best known examples of fake news was the story attributed to the ‘Denver Guardian’ (no such newspaper exists), which claimed that the FBI agent suspected of leaking the emails from Hillary Clinton during the 2016 US election campaign had been found dead, murdered in a way that made it look like suicide. At one point this item was being shared on Facebook at one hundred times per minute, generating hundreds of thousands of views.

The history of fake news

Although the term gained notoriety during the 2016 US presidential election campaign, the phenomenon as such is much older: it is as old as mass communication. Purposely written, false stories—for instance, Procopius’ Anecdota—existed in the Roman Empire and were abundant in eighteenth-century France and England. Partly or fully false stories also played an important part in the success of Joseph Pulitzer’s and William Hearst’s ‘yellow press’ at the end of the nineteenth century. In the 1930s the League of Nations carried out research into how to combat ‘false news’ most effectively. While versions of fake news have thus long been present in society, the fake news that exists in the realm of online social media is different, not least because of the speed at which it can be propagated and the reach it can attain.

---

20 Gillin, ‘Fact-Checking Fake News’.
22 Ibid.
Trapped in bubbles

Social media bubbles (also known as echo chambers) are yet another alarming feature of social media, especially Facebook. ‘Bubbles’ are groups of users who consume—whether consciously or not—the same content and are basically not offered alternative information or opinions. Bubbles are generated automatically: based on the previous online behaviour of a given user, social media algorithms decide on the content that is shown to the user and that which is buried deep in the newsfeed (making it effectively non-existent). According to a UN report on the promotion and protection of the right to freedom of opinion and expression, ‘[p]latforms deploy algorithmic predictions of user preferences and consequently guide the advertisements individuals might see, how their social media feeds are arranged and the order in which search results appear’.²⁵ The number of possible bubbles is infinite, but the general logic of bubble creation seems to be simple: extreme right-wing militants will see extreme right-wing content; likewise, radical left-wing users will be surrounded by content posted by like-minded users; and so on. However, the details of how the newsfeed algorithms operate remain unclear, as noted by Philippe Laloux from *Le Soir*:

I’m the editor of *Le Soir*’s Facebook profile; I oversee our newspaper’s online activity. I use Facebook daily, for professional purposes. And you know what? I can’t see any of *Le Soir*’s stories on my private Facebook profile newsfeed, the one I use to chat with friends. Why that is, I cannot fully understand. Both Facebook and Google are like black boxes.²⁶

The automated process of bubble creation is made easier by human psychology. As Nobel Prize winner in economics Daniel Kahneman has explained, humans show an inherited tendency to ignore facts that force their brains to work harder²⁷—which is exactly why we like to stay within the bubble. ‘Humans tend to cluster among like-minded people and seek news that confirms their biases.

²⁶ P. Laloux, interview.
Facebook’s research shows that the company’s algorithm encourages this by somewhat prioritizing updates that users find comforting’, notes Zeynep Tufekci, a leading social media researcher.28 As individuals tend to prefer to receive information confirming their pre-existing views, information coming from unfamiliar sources and/or contradicting their pre-existing views tends to be ignored. ‘As a result, correcting misinformation does not necessarily change people’s beliefs’, according to research from Harvard Kennedy School.29

As comforting as they are, such bubbles may lead to a serious distortion of the public debate. Jürgen Habermas warned that in liberal regimes, online media, with their millions of isolated forums for debate on a vast range of topics, could lead to a ‘fragmentation of the public’ and a ‘liquefaction of politics’.30 In a recent report, Harvard Kennedy School researchers warned that the end result of persistent echo chambers would be a lack of shared reality, which would clearly be dangerous to democratic society.31

Bots: social media machine guns

Another worrying phenomenon seen on social media platforms, this time mostly on Twitter, is the existence of automatic profiles, or ‘bots’: special programs able to operate autonomously, mostly pretending to be a genuine profile managed by a human being.32 Bots can mass-send content, and share or retweet chosen items at machine-gun speed, reacting to a pre-programmed sequence of words, hashtags and so on. They can also follow each other, creating the pretence of a popular profile.33 Bots’ main goal is to fill social media with messages, more or less targeting pre-programmed groups of re-

31 Lazer et al., Combating Fake News.
33 Ibid.
 Much of the fake news that can be seen on social media is spread by bots. In 2015 no less than 5% of all comments posted on Twitter were generated by bots. Roughly 20% of the tweets sent during the 2016 US presidential campaign were artificially generated.

---

How fake news and online bubbles impact politics
The 2016 US presidential elections illustrate why the spread of fake news is potentially dangerous. In the months preceding the vote, fake stories related to the elections generated more engagement (shares, likes and comments) than the 20 best-performing real news stories. The 20 most popular fake stories from hoax sites and partisan-biased blogs scored 8,711,000 reactions on Facebook. The 20 best-performing true stories from 19 major news websites generated only 7,367,000 reactions.\(^{37}\) The sheer ubiquity of disinformation must have left a mark. One particular piece of fake news—the story about Pope Francis allegedly endorsing Donald Trump—was shared millions of times.\(^ {38}\) Bearing in mind that there are approximately 70 million Catholics living in the US, for whom Pope Francis is the unquestionable spiritual leader, the effect of such a story may have been significant.

The US presidential campaign was also exceptional because the leading candidate peddled fake news himself. According to the press, on at least 10 occasions Donald Trump published false content on his Twitter account, including claims that President Barack Obama’s birth certificate was fraudulent,\(^ {39}\) that official unemployment data was fabricated and that illegal immigrants were being allowed to vote on a huge scale. It remains unclear whether he was aware that the information he shared was fake.\(^ {40}\)

EU countries have been affected by the spread of misinformation too. Take the example of the referendum campaign in Britain in 2016. One of the most resonant arguments of the pro-Brexit camp was that UK taxpayers pay £350 million to the EU each week. ‘It was clearly the most effective argument not only with the crucial swing fifth [of the voting population] but with almost every demographic…. Would we have won without [the claim that £350 million could fully fund the National Health Service]? All our research and the close result strongly suggests No’, admitted Dominic Cummings, Director of the Leave Campaign.\(^ {41}\) Nobody seemed to care that it was a false claim: as early as April 2016 the UK Statistics

---


\(^{38}\) Z. Tufekci, ‘Mark Zuckerberg’.


\(^{40}\) Ibid.

\(^{41}\) D. Cummings, ‘Dominic Cummings: How the Brexit Referendum was Won’, *The Spectator*, 9 January 2017.
Authority, the official body policing the use of statistics, had dismissed the £350 million figure as inac-
curate.42

Some of the fake news that has surfaced in the recent past in both Europe and the US could have
been created by agents of foreign powers, with the aim of destabilising Western countries and fuelling
social tensions. On 6 January 2017, the US Office of the Director of National Intelligence published a
report on Russia’s role in influencing the US presidential elections. It openly stated that there had been
a Russian-originated, propaganda-cum-disinformation campaign. It was aimed not only at disturbing
the electoral process, but also at helping to elect Donald Trump as the next president.43 In 2016 more
than 200 websites peddled Russian propaganda, reaching an estimated American audience of 15 mil-
lion people. On Facebook alone, fake news stories planted by the Russian propaganda machine were
viewed more than 213 million times.44 It is not possible to determine whether the Russian campaign
proved decisive in electing Trump, but it at least proved rather effective at sowing distrust of US democ-
racy and its leaders.45

In February 2016, General Petr Pavel, Chairman of the NATO Military Committee, accused Russia of
spreading a false report claiming that a German soldier stationed in Lithuania had raped a local girl (in
reality, nothing of the sort had happened). In Pavel’s view Russia is trying to undermine the support for
NATO troops deployed in Eastern flank countries.46 The famous ‘Lisa case’, dating back to the beginning
of 2016, is yet another example of possible foreign interference. The story, which circulated in Germany,
claimed that a girl of Russian origin named Lisa had been raped by refugees from the Middle East. The
story was untrue but went viral, causing widespread outrage against refugees (and prompting official
condemnation from Russia). In 2017, as the parliamentary election date approaches, the German press
has reported on the circulation of a rising number of misleading and outright fake stories about Chan-

---

45 Ibid.
cellor Angela Merkel.\textsuperscript{47} Germany’s domestic security agency has warned against cyber-operations that could potentially endanger German government officials, members of parliament and the employees of democratic parties.\textsuperscript{48}

Similar patterns appear in many other European countries. Ukraine was probably the first country to be targeted by Russian-originated disinformation on a mass scale. A flood of anti-Ukrainian fake news engulfed the country in spring 2014, when Russia’s ‘green men’ invaded Crimea. More recently, in April 2017, the Dutch intelligence service reported that Russia had tried to influence the results of the March 2017 Dutch parliamentary elections by spreading false information ‘to push voters in the wrong direction’.\textsuperscript{49} Sources linked to Russia were also reported to be very active during the French 2017 presidential elections, peddling dubious polling data and falsely accusing Emmanuel Macron, then the leading candidate, of many different misdeeds.\textsuperscript{50} NiemanLab reported that certain fake news sites had achieved an astonishing level of sophistication: one site that mimicked the Belgian site \textit{Le Soir} even included links that led back to the real \textit{Le Soir} website.\textsuperscript{51} Fortunately for Macron, efforts to smear him were met with stiff resistance: various teams of experts identified and countered harmful disinformation.\textsuperscript{52} In Italy, a journalism investigation revealed the existence of a whole network of highly popular social media profiles spreading anti-EU and anti-US propaganda.\textsuperscript{53}

Some fake news is generated for financial purposes. The mechanism used to be simple: perpetrators would create websites and publish fake content there. In parallel, they would join online advertising schemes (e.g. Google Ads) and start posting links to their stories on Facebook or Twitter, to attract social media users. The more traffic was directed to their sites, the more money they would earn from advertising. Some entrepreneurial spirits from Veles, a town in the centre of the former Yugoslav Republic of

\begin{itemize}
\item \textsuperscript{47} A. Nardelli and C. Silverman, ‘Hyperpartisan Sites and Facebook Pages are Publishing False Stories and Conspiracy Theories about Angela Merkel’, \textit{BuzzFeed}, 14 January 2017.
\item \textsuperscript{51} S. Wang, ‘The French Election is Over. What’s Next for the Google—and Facebook—Backed Fact-Checking Effort There?’ \textit{NiemanLab}, 8 May 2017.
\item \textsuperscript{52} A. Alaphilippe and N. Vanderbiest, Saper Vedere consultancy, interview with author, Brussels, 29 May 2017.
\item \textsuperscript{53} A. Nardelli and C. Silverman, ‘Italy’s Most Popular Political Party is Leading Europe in Fake News and Kremlin Propaganda’, \textit{BuzzFeed News}, 29 November 2016.
\end{itemize}
Macedonia with a population of just 50,000, discovered that flashy stories on Hillary Clinton attracted a lot of interest—so they produced them in big quantities.54

The phenomenon of fake news on social media being relatively new, only a limited amount of scientific research has been carried out on it. Scholars agree that fake content has a potentially negative effect on political life,55 but so far this effect has not been measured.56 Some researchers point to the fact that not all Internet users are exposed to fake news: a survey conducted in the US in December 2016 reported that only 33% of those polled recalled seeing fake news headlines57 (although it is not clear whether those surveyed were objectively capable of distinguishing the false from the true). Having said that, the same survey showed that the majority of those who had actually seen fake headlines considered them to be either ‘very accurate’ or ‘somewhat accurate’, which indicates the potential of fake news to deceive.57

Media practitioners are less cautious. They openly say that fake news is a real danger to democracy, as political choices should be based on facts and informed opinion, not lies.58 According to David Schraven, founder of the German fact-checking website, Correctiv, ‘[t]his is indispensable in a democracy. We are currently seeing this process being interrupted by many targeted false logs in social networks’.59 When people start to believe fake news60 instead of genuine information, it corrupts the credibility of democratic society’s system for information sharing. The lack of trusted bearings undermines the very structure of society.61

Facebook’s own research confirms that the content it publishes could influence the outcome of political processes. As early as 2010, the company conducted an experiment on 61 million users in the US: it

---

55 D. Wolter, interview.
57 Ibid.
59 Ibid.
60 Davies, ‘Interview with Craig Silverman’.
divided them into two groups, showing each of them different types of posts inciting them to vote in the mid-term elections. The company then checked the public voter rolls to verify which of the two groups had a higher voter turnout. The differences were marked.\footnote{Tufekci, ‘Mark Zuckerberg’.}

On at least one occasion, fake news has led to a physically dangerous situation. In October 2016, a fake story, claiming that a paedophile ring linked to Hillary Clinton was being run from one of Washington, DC’s restaurants, emerged on social media. Despite being debunked, the story kept resurfacing. Then, on 5 December 2016, 28-year-old Edgar M. Welch, armed with an automatic rifle, stormed into the restaurant and started to shoot. Fortunately, nobody was hurt. When arrested, Welch explained that he wanted to ‘investigate’ what he had earlier read on one of the social media platforms.\footnote{E. Lipton, ‘Man Motivated by “Pizzagate” Conspiracy Theory Arrested in Washington Gunfire’, \emph{The New York Times}, 5 December 2016.}

Therefore, it is no wonder that fake news gives rise to concern. A study conducted in Canada and published in February 2017 showed that 52\% of the Canadians polled were ‘worried’ about the impact of fake news and a further 29\% were ‘somewhat worried’, with only 18\% declaring otherwise.\footnote{S. Krashinsky-Robertson, ‘Fretting Over Fake News’, \emph{The Globe and Mail}, 21 February 2017.} An earlier poll conducted in the US showed similar results: as many as 88\% of those surveyed agreed that false news caused ‘a great deal of’ or at least ‘some’ confusion.\footnote{L. H. Owen, ‘How to Cover Pols Who Lie, and Why Facts Don’t Always Change Their Minds: Updates from the Fake-News World’, \emph{NiemanLab}, 21 February 2017.}

Social media bubbles add to the distortion of reality. They reinforce emotions. The stronger the echo chambers are, the more polarised societies become. This is what populists from all sides of the political spectrum crave: a polarised society creates the space to develop an ‘us versus them’ logic.\footnote{J. A. Emmanouilidis and F. Zuleeg, \emph{EU@60—Countering a Regressive & Illiberal Europe}, European Policy Centre (Brussels, October 2016), 25.} The prevalence of bots is no less concerning. One can theoretically control infinite numbers of bots, which creates the ability to influence—if not outright manipulate—the physics of social media, namely the ‘trends’ on social networks. An army of bots can give the impression that one particular political opinion is more popular than another. This adds ‘fake popularity’ to ‘fake news’. One botnet discovered in Ukraine was composed of roughly 15,000 automated Twitter profiles. Of all the Twitter accounts following both Hillary

\footnote{Tufekci, ‘Mark Zuckerberg’.}
Clinton and Donald Trump during the presidential election campaign, 40% were bots.67 A third of all Twitter traffic related to the 2016 UK referendum campaign (Brexit) was probably caused by bots.68 The massive presence of bots may even have endangered the integrity of the 2016 presidential election.69 And as with the creation of fake news, the influence of an ill-intentioned foreign power might be at play. In March 2016 The Economist warned that authoritarian governments were investing heavily in web-based propaganda infrastructure.70

The gravity of the problem has been recognised by political leaders. In the wake of the US presidential election campaign outgoing president, Barack Obama admitted that ‘if we are not serious about facts and what’s true and what’s not, if we can’t discriminate between serious arguments and propaganda, then we have problems’.71 Likewise, one month earlier German Chancellor Angela Merkel, in a speech at a conference in Munich, had warned that search engines and social media distort people’s perceptions of reality, which in turn negatively impacts democracy.72 In November 2016 the European Parliament adopted a resolution on ‘EU strategic communication to counteract anti-EU propaganda by third parties’. It called on EU member states to adapt and update legislation to address ongoing developments.73

Other voices followed, such as that of Italian Competition Authority President Giovanni Pitruzzella. Democracies, he said, faced a simple choice: leave the online realm as it is, as a sort of ‘Wild West’, or implement regulations that reflect the changed context of communication. He called for the creation of independent authorities, coordinated by Brussels and able to impose fines and remove fake news, arguing that it is the job of public powers to ensure the veracity of information.74

68 Miller, ‘Governments Don’t Set the Political Agenda Anymore’.
69 Bessi and Ferrara, ‘Social Bots Distort’.
Even the official lobby of the British press, the News Media Association, has called on national authorities to take action against the spread of fake news, as it causes real harm by facilitating the spread of conspiracy theories.\textsuperscript{75}

The current regulatory framework for social media
Social media in Western countries operate in a specific environment of ‘legal exceptionalism’, as The Economist puts it.\(^7^6\) With some narrow exceptions, companies are not responsible for the content published on their platforms. This unique legal context is the result of a particular mindset, originating 20 years ago, shared by the US and the EU: a combination of a genuine willingness to protect the (then) fledgling industry and an apparent lack of understanding of the potential of social media platforms.

In 1996 the US Congress amended the Communications Decency Act, granting companies immunity from prosecution for the content posted by third parties. The same legal logic was applied in the 1998 US Digital Millennium Copyright Act. Exemptions from the immunity were rare, being limited to copyrighted material and serious crime (such as child pornography or incitements to violence).

The EU followed suit. The 2000 E-Commerce Directive defined social media platforms as ‘Internet service providers’,\(^7^7\) not liable for stored content as long they are not aware of its illegal character and handle it passively, without modifying it. The latter is crucial: the liability protection applies only when the platform’s role is limited to the technical process of transmitting and temporarily storing the content. Should the content be clearly illegal, such as child pornography, both the American and the European legal systems provide a way to remove it. This is called the ‘notice and take down procedure’. The US Digital Copyright Millennium Act 1998 laid out the specific form of a ‘notice and take down procedure’.\(^7^8\) The European system left the details to national legislation and authorities. However, in both the EU and the US it is a third party—public authorities, for example—that notifies the social media platform of the illegality of its content. Companies are not obliged to discover it for themselves. The E-Commerce Directive explicitly states that the ‘providers’ shall not be obliged to monitor the information which they transmit or store. In May 2016, on the initiative of the European Commission, the leading social media brands—Facebook, Twitter, YouTube and Microsoft—agreed a voluntary code of conduct. Its aim was to improve ways of combating hate speech and to make the ‘notice and take down procedure’ more effective. The companies obliged themselves to remove illegal content (based

\(^7^6\) The Economist, ‘Eroding Exceptionalism’, 11 February 2017, 49.


\(^7^8\) Article 19, Internet Intermediaries, 7.
on a review by online intermediaries) within 24 hours of receipt of notification of its existence. The agreed code of conduct also obliged social media giants to have their own guidelines to clarify to users what constitutes illegal hate speech.79

Lies, however harmful, are not considered ‘illegal content’. Some have learned this the hard way. In 2015 Anas Modamani, a 19-year-old refugee from Syria living in Germany, had the chance to be photographed with Chancellor Angela Merkel, who was visiting the refugee shelter where he was staying. The very amiable and symbolic picture hit the front pages. Unfortunately, a few days later the same picture started to appear in fake stories about terrorist threats in Europe. Some of them, distributed mainly via Facebook, linked Modamani to the attacks in Brussels and Berlin. Many of these stories are still online, including a Facebook post by a certain K. F. which has been shared almost one thousand times.80 In total, stories alleging or suggesting that Mr Modamani was a terrorist generated more than 32,000 shares, reactions and comments on Facebook. By contrast, the Deutsche Welle story debunking these lies generated only 13,000 shares and reactions.81 Faced with digital defamation, in February 2017 Modamani sought a court injunction against Facebook. He and his lawyer wanted the court to order Facebook to stop the reposting of the picture in stories linking him to terrorism. Yet in March 2017, the court in Würzburg ruled that that there were no grounds in existing law for such an action. The court also declared that Facebook could only be held responsible for deleting the disputed content if it was considered technically possible—and Facebook claimed it was not.82

There are other individuals who have tried to fight fake news on their own. Austria’s Green Party (Die Grünen) leader, Eva Glawischnig, filed a complaint with Vienna’s commercial court, asking for the immediate removal from Facebook of a fake news story concerning her. Her legal team followed this up with an injunction against Facebook’s European headquarters in Dublin.83 At the time of writing, the case had not been resolved. However, regardless of the final verdict, it is hard to compare the chances

80 The identity of the person posting this outrageous claim against Modamani has been omitted at the author’s discretion. The print screen and web address of this post have been archived.
81 Nardelli and Silverman, ‘Hyperpartisan Sites’.
of an established political figure, supported by a party apparatus, to those of an individual person, such as Mr Modamani. All in all, it is fair to conclude that new channels of online content distribution have been developing faster than society’s ability to understand and frame them.⁸⁴ ‘Facebook’s power and the lack of accountability of that power is a problem in itself that needs to be addressed’, claims Ben Wagner, Stiftung Wissenschaft und Politik associate and researcher on digital policies.⁸⁵

⁸⁴ Lazer et al., Combating Fake News.
⁸⁵ B. Wagner, interview.
The industry’s voluntary measures to tame fake news
Flagging, reporting and debunking

By the end of 2016 the criticisms of social media’s apparent indifference to fake news had started to get louder. The opinion-forming press argued that the companies owning and managing the social media platforms, such as Facebook and Google, bear at least some responsibility for the flood of fake news, and that therefore they need to engage in checking the further spread of the virus.86

Echoing these calls, the social media platforms’ representatives acknowledged the problem. On 15 December 2016, Facebook announced its first measures to address the issue. The company tackled the most blatant cases of fake news and hoaxes spread by spammers for financial gain. The approach Facebook chose to use to fight fake news placed the burden of responsibility onto users. It is up to users to report alleged fake news by clicking on a special dialogue window that has been added to the platform’s interface. Once reported by a certain number of users, the contested content is then passed on to independent fact-checkers. If it is confirmed as fake, it is flagged as ‘disputed’. It can be further shared, but users are warned about the controversy surrounding it. This new feature was first implemented in the US: in March 2017 newspapers reported that the first items carrying the glowing red alert signs had been spotted.87 ‘We believe in giving people a voice,’ explained company Vice-President Adam Mosseri,88 emphasising that the company was not intending to become the ‘Ministry of Truth’. In mid-January 2017 Facebook announced that the flagging and fact-checking tools would also become available in Germany, as a pilot project. Correctiv was initially chosen as the third-party, independent fact-checker.89 However, a month later, Facebook put the test phase on hold, as it had been unable to find more partners.90 Mathias Döpfner, Chief Executive Officer (CEO) of Axel Springer, one of the biggest German media conglomerates, had ruled out helping Facebook. In his view, it would be a mistake for publishers or public broad-

casters to help social media solve their credibility problem.\textsuperscript{91} The third and fourth countries chosen to be testing grounds for the anti-fake news measures were the Netherlands and France, respectively. In contrast to their German peers, the leading French media companies teamed up with Facebook in this endeavour.\textsuperscript{92} The choice of countries was not random: in 2017 the Netherlands, France and Germany are all holding elections. The Facebook approach, of relying on users’ engagement to spot fake news, has been applauded by commentators, who believe that empowering the community is the best way to police social media.\textsuperscript{93}

In April 2017 Facebook followed up with a paper on ‘Information Operations’, in which it elaborated on its intention to fight the fake news and illegal content posted online. The added value of Facebook’s paper, co-authored by the company’s senior security officers, is that it details the company’s view on the definition of fake news and related problems. For example, Facebook distinguishes ‘disinformation’ (inaccurate content spread intentionally) from ‘misinformation’ (the unintentional spread of inaccurate information). It also identifies ‘false news’ as items that ‘purport to be factual, but which contain intentional misstatements with the intention to arouse passion, attract viewership or deceive’.\textsuperscript{94} Moreover, Facebook openly acknowledged the existence of ‘information operations’: actions taken by organised actors (both governments and non-state agents) aimed at distorting political sentiments, most frequently to achieve a strategic outcome.\textsuperscript{95} According to the company’s research, there are three common objectives of ‘information operations’: promoting or denigrating a specific issue, sowing distrust in political institutions and spreading confusion. Facebook’s paper also acknowledged the fact that false news can be amplified not only by the creators of the disinformation—through their networks of false accounts and so on—but by everyday users as well, unintentionally, through authentic networks. The company recognises that the motivation of the false amplifiers is ideological rather than financial. The paper also offered additional insights into how the company intends to fight malicious ‘information operations’. For example, it underlined the company’s resolve to prevent and delete fake accounts, whether they are manually or automatically

\textsuperscript{93} K. Iszkowski, ‘Granice wolności słowa w social media’ [Limits to the Freedom of Speech in Social Media], \textit{Libertél}, 1 December 2016.
\textsuperscript{94} J. Weedon, W. Nuland and A. Stamos, \textit{Information Operations and Facebook}, Facebook (27 April 2017), 5.
\textsuperscript{95} Ibid, 4.
operated. The company also claimed that in France alone, in the wake of the presidential elections, it had suppressed 30,000 fake accounts.\textsuperscript{96}

Other methods to stop fake news are employed too. Snap, the owner of the Snapchat platform, has asked third parties publishing on its ‘Discover’ platform to vouch for the content they provide. As the company’s spokesman explained, it wants its editorial partners ‘to do their part to keep Snapchat an informative, factual, and safe environment.’\textsuperscript{97} Another way to weed out some of the fake news, suggested by experts, is to enforce the existing real-name policy more strictly (in theory, users are obliged to publish under their real names or at least to register with platforms with genuine individual data).\textsuperscript{98} In the near future, social media platforms might also be willing to use artificial intelligence to comb through the content even more efficiently. In February 2017 it was announced that Google had launched the first artificial intelligence tool able to find abusive comments without human assistance. ‘The Perspective’, as this software is called, does not delete abusive content on its own but reports the questionable content to human editors, allowing them to decide whether the given item should (or should not) be taken down.\textsuperscript{99}

There are other bodies which, independent of social media platforms, try to limit the plague of fake news. At the forefront are fact-checking establishments such as the US-based FactCheck.org, Snopes.com and Truthoorfiction.com. The ‘old’ media have joined this effort too. The French newspaper \textit{Le Monde} is one such example: it has established a special unit called ‘Les Décodeurs’, which not only fact-checks popular stories, but also offers access to ‘Decodex’—a special database of 600+ websites identified as being sources of fake news.

Some initiatives specialise in fighting the fake news generated by foreign powers. StopFake.org, a project created by Ukrainian media workers and academia members to debunk Russian propaganda, is one of them.\textsuperscript{100} Another is East Stratcom, a special unit of the European External Action Service, which also focuses on debunking Russian propaganda. During its 16 months of existence, it has exposed 2,500

\textsuperscript{96} Ibid.
\textsuperscript{98} The real-name policy means that the social media user is obliged to use his or her real name, the one written on the given person’s identification. See: Facebook, ‘Statement of Rights and Responsibilities’, last updated 30 January 2015, art. 4.
false stories.\textsuperscript{101} However impressive, the number is still a drop in the ocean. It remains to be seen whether fact-checking outlets such as Decodex will become truly popular. It is possible that only a handful of power-users, those who are less likely to fall for fake news anyway, will subscribe to it.\textsuperscript{102}

\textbf{The limits of self-imposed measures}

Past experience suggests that weeding out unwelcome content is not an easy task. According to an initial evaluation of the application of the ‘code of conduct’, Facebook removed only 28.3\% of illegal content within the set timeframe of 24 hours (the figure for Twitter was 19.1\% and for YouTube was 48.5\%).\textsuperscript{103} A similar lack of effectiveness was observed regarding the commitments made previously (in 2015) by social media platforms in Germany:\textsuperscript{104} as little as 46\% of the notified content was deleted.\textsuperscript{105}

The industry believes that it can improve its track record. As mentioned earlier, some social media giants have experimented with automated content filtering. Yet the use of artificial intelligence raises serious doubts. What if its algorithm is badly designed? What if it discriminates? How can the transparency of the process be guaranteed?\textsuperscript{106}

It is far from certain whether Facebook’s preferred way to combat fake news—fact-checking and flagging instead of removing—is good enough. It offers advantages but it might be too weak to stop the spread of false information.\textsuperscript{107} Moreover, the mere act of fact-checking is meaningless for those who consume

\textsuperscript{103} European Commission, Directorate-General for Justice and Consumers, ‘Code of Conduct on Countering Illegal Hate Speech Online: First Results on Implementation’, Factsheet, December 2016.
\textsuperscript{104} P. Oltermann, ‘Germany to Force Facebook, Google and Twitter to Act on Hate Speech’, \textit{The Guardian}, 17 December 2016.
\textsuperscript{105} Bundesministerium der Justiz und für Verbraucherschutz, ‘Löschung rechtswidriger Hassbeiträge bei Facebook, YouTube und Twitter’, 26 September 2016.
\textsuperscript{107} S. Wang, ‘A Threat to Society.
fake news in order to confirm their existing biases or for readers who distrust traditional media. Facts are too often subconsciously disregarded if they contradict a deep conviction. Flagged or not, such users would probably still read the item. Besides, the measures implemented by Facebook might prove to be a double-edged sword. Just imagine a group of ill-intentioned users reporting genuine information and verified stories as ‘fake’. If done on a massive scale, it could clog the system, whether formed of automated algorithms or human fact-checkers.

Historical experience of the success of self-imposed measures in other areas is not encouraging either. Take roaming charges, for instance: few remember that the problem of the excessive cost of mobile phone roaming was raised as early as 1999. Back then the European Commission opted for the hands-off approach, expecting the telecom operators to solve the problem. Years passed and virtually nothing changed: in the mid-2000s it was realised that the reductions in wholesale costs were hardly ever being passed on to the retail customer. Eventually, the Commission and the European Parliament decided to act. The rest is history: subsequent legislative acts, starting with the 2007 Regulation, cut the prices. Ten years later another piece of legislation has abolished roaming charges altogether. According to Wired magazine, ‘not only is self-regulation largely a fantasy, but repeated scandals across multiple industries have proved that companies are fundamentally incapable of self-regulating for the greater good’.

108 Ibid.
109 Owen, ‘How to Cover Pols Who Lie’.
A new approach to the problem of fake news
As the case of roaming charges shows, state-imposed legislation is sometimes necessary to protect the interests of the public. It would become easier to imagine such a step if social media platforms started to be considered as fully fledged media companies, like newspapers, TV channels or radio stations.

There are reasons to believe that social media are no longer ‘Internet intermediaries’. It can be argued that Facebook actively manages content—the ‘trending’ section is the best example of this.\(^\text{113}\) In 2016, the company acknowledged that it uses human editors to pick and evaluate trending topics, thus meeting the Council of Europe’s definition of an editorial process.\(^\text{114}\) Robert Thomson, CEO of News Corp., has claimed that social media platforms can no longer be considered intermediaries only:

> These companies are in digital denial. Of course they are publishers and being a publisher comes with the responsibility to protect and project the provenance of news. The great papers have grappled with that sacred burden over decades and centuries, and you can’t absolve yourself from that burden or the costs of compliance by saying, ‘We are a technology company’.\(^\text{115}\)

Martin Sorell, CEO of WPP, the world’s largest advertising company, has also argued that social media should be responsible for the content in their ‘digital pipes’.\(^\text{116}\)

In December 2016 Mark Zuckerberg, Facebook’s CEO, himself admitted that the company is a ‘media company’, and not just a mere ‘platform’.\(^\text{117}\) Zuckerberg’s admission is corroborated by the newest features to have been added to Facebook, such as ‘Facebook Live’ (a functionality allowing users to broadcast a live video stream), and the company’s decisions to buy video content from external media companies and to introduce its own TV-style shows. Today’s social media can hardly be compared to the service provided by telecom companies, which strictly limit their activity to offering physical access to the network. The very fact that there is a complicated (and automated) mechanism to filter,

---

\(^1\) To clarify, ‘trending topics’ are a separate part of the Facebook platform. They are selected, according to Facebook, because they merit users’ attention due to their subjectively assessed importance and popularity among other users.

\(^2\) N. Helberger and D. Trilling, ‘Facebook is a News Editor: The Real Issues to be Concerned About’, Media Policy Project blog, 26 May 2016.

\(^3\) Waters, Garrahan and Bradshaw, ‘Harsh Truths’.

\(^4\) Ibid.

\(^5\) M. Ingram, ‘Mark Zuckerberg Finally Admits Facebook is a Media Company’, Fortune, 23 December 2016.
categorise and structure the incoming content makes the difference. In addition, Facebook, Twitter and Google can no longer be considered start-ups in need of special protection: they are multibillion dollar companies. In 2016 alone Facebook reported $27.6 billion in total revenues and a whopping net profit of $10.2 billion.\(^{118}\)

**Top-down regulation imposed so far**

As already mentioned, state-imposed regulations have not provided the tools needed to fight fake news so far. Only the presence of illegal content has been addressed. At the time of writing, the German government is leading the way in this. In March 2017 it tabled a proposal to impose hefty fines (of up to €50 million) on social media companies that fail to take down illegal content quickly enough. The draft law would give social media platforms 24 hours to delete the most blatant cases.\(^{119}\) For less offensive content, the time allowed would be extended to one week—with the clock starting once the first complaint is filed. According to German Justice Minister Heiko Maas, Facebook and Twitter have missed the chance to improve their take-down practices themselves.\(^{120}\) Other European governments—including the British government—have acted more cautiously, limiting themselves to calls for additional voluntary measures.

But this may change. Following the deadly terror attack in London, the British government stated that ‘companies like Facebook and Google can and must do more’ to remove potentially harmful material.\(^{121}\)

---

\(^{118}\) Facebook, ‘Facebook Reports Fourth Quarter and Full Year 2016 Results’, Press Release, 1 February 2017.


\(^{120}\) Ibid.

\(^{121}\) A. Sparrow, ‘Social Media Companies such as Facebook and Google “Can and Must Do More” to Remove Radical Material—PM’s Spokesman’, *The Guardian*, 24 March 2017.
A new proposal for the short and medium term: the ‘notice and correct’ procedure

Based on the assumption that social media are in fact owned by media companies, this paper calls on governments to consider applying a single, real-life-tested and effective tool to combat fake news: the existing press laws. Enforced in most of the EU member states, they create a legal obligation to correct false information. To some extent, they also play a preventative role: the media have learned to take precautionary measures, knowing that they bear responsibility.

If press laws were to be applied, social media platforms, like the traditional press, would have to correct (or take down) false information at the request of the genuinely affected party. Should the platform decide to ignore the request (which it would be entitled to do), the affected party would have the right to refer the case to an independent court, exactly as is the case with newspapers in most EU countries. The same right of referral to the courts should be guaranteed to the authors of the removed content, allowing them to try to prove that there is nothing wrong with the contested item.

The procedure of correction should be similar (in its results) to the one applied to print publications. Any party notifying the platform should receive an obligatory confirmation of receipt. Once notified, the company (platform) would have to choose one of the following alternative courses of action: (1) do nothing and face the possible eventuality of being referred to the court; (2) fact-check the item concerned and then decide whether it should stay, be corrected or be deleted; (3) delete the item outright, based on an in-house assessment (for example, by the in-house fact-checking team); or (4) ask the author of the content, if identifiable, to correct the information, under threat of deletion if it is not corrected.

In this framework, the contested item could be corrected either by the author(s) or by the social media platform itself. It is important to retain this option. As already explained, many fake news stories are created intentionally, by agents that cannot be coerced into deleting them (because they are the functionaries of a foreign country’s secret service apparatus, for example). In such a case, the platform
must take responsibility and intervene without the authors’ consent. Depending on the platforms’ technological capacities, an additional element to the procedure might be considered as well: an obligation to distribute the corrected content to exactly the same audience as that which read the fake story in the first place. Such an obligation would mimic the obligations imposed in some countries on the printed press, which require it to publish the correction in the same medium, in the same place as the false information it has corrected.\textsuperscript{122}

This paper proposes that the whole course of action, as described above, is called the ‘notice and correct’ procedure—to underline its different character (and different legal basis) from the ‘notice and take down’ procedure, which is related to content identified as illegal in the light of penal codes.

In order to allow such a change, current legislation, such as the E-Commerce Directive, would have to be rewritten. The legal logic should be reversed: in principle, social media platforms would become legally responsible for the content published, just as the traditional media are. In this softened version, the legal responsibility could kick in at the moment that the platform in question receives notification of the content’s fake nature. However, one thing should not change: the companies would not be obliged to monitor all the content published and stored on their platforms or to proactively search for fake content—such a demand could prove to be too much of a burden. As already noted, the ‘notice and correct’ procedure would be triggered at the request of a genuinely affected party.

The proposed procedure should be imposed after a waiting period. Although the author of this paper firmly believes that current and past experiences justify important regulatory changes, the benefit of the doubt should be given to the owners of the platforms, who claim that self-imposed measures will be enough to address the problem of fake news. Submitting this claim to a test, over a limited period of time (one to two years) would ultimately settle the issue.

\textsuperscript{122} One of these countries is Belgium. See Belgium, Loi du 23 juin 1961 relative au droit de réponse, Belgisch Staatsbal Moniteur Belge, 9 February 2010, 7844.
What are the risks attached to the use of press laws?

The eventual introduction of the notice and correct procedure needs to be weighed in the context of freedom of speech. In both the EU and the US freedom of speech is considered a fundamental value and is protected by primary legal acts, such as the EU member states’ national constitutions, the European Charter of Fundamental Rights (art. 11)\(^{123}\) and international treaties. The International Covenant on Civil and Political Rights allows very limited scope for potential restrictions to free speech: only to respect the rights or reputations of others and for the protection of public order, security, public health and morals.\(^{124}\) As Mr Boni underlined, ‘any form of censorship should be prohibited. The freedom of expression is one of the key values in our civilisation. At the same time it is important to design a new framework for the functioning of digital platforms’.\(^{125}\)

Real-life-tested press codes would enable this issue to be addressed properly. After all, freedom of speech is not limitless. It is enjoyed only within some sort of framing, such as ‘enhancing the access to and the diversity and quality of the channels and the content of communication’.\(^{126}\) It would also be wrong to debate its limits separately from the current context.\(^{127}\) The goal of most haters online is to stifle the discussion, to silence those who dare to have opposing views, or who support a different political party or take a different position on a given social/economic issue. It would be rather naïve to guarantee totally unrestricted freedom of speech to those whose long-term aim is to destroy democracy and its freedoms altogether. Unrestricted free speech, devoid of any form of commonly shared rules,

---

\(^{123}\) EU, Charter of Fundamental Rights of the European Union, OJ C326 (26 October 2012), 391.

\(^{124}\) Article 19, Internet Intermediaries, 10.

\(^{125}\) M. Boni, interview.


\(^{127}\) M. Wojciechowski, ‘Granice wolności słowa w social media’ [Limits of Free Speech in Social Media], Libertél, 1 December 2016.
would be the equivalent of a snake eating its own tail.\textsuperscript{128} As German Minister of Justice Heiko Maas has argued, defamation and malicious gossip should not profit from the freedom of speech principle.\textsuperscript{129}

Western democracies have been debating this issue for many years and seem to have found a delicate balance between free speech and its restrictions, justified by the public interest.\textsuperscript{130} The old media, that is, newspapers, radio and TV, have been regulated for decades, if not centuries (as is the case for the printed press). For example, in the Netherlands sanctions for publishing false information date back to the seventeenth century,\textsuperscript{131} in Denmark press regulations were introduced in the mid-1800s,\textsuperscript{132} and in Finland the press code is 100 years old.\textsuperscript{133} Belgian federal law contains a number of restrictions based on the need to protect the public interest, such as those on matters relating to national security, moral standards and the monarchy.\textsuperscript{134} In the Czech Republic, the Press Act of 2000 stipulates that the publisher is legally responsible for the trustworthiness of published information,\textsuperscript{135} as does the Italian press law of 1981, which regulates the penal responsibility of both publishers and journalists.\textsuperscript{136} In Germany, the press is regulated at \textit{Länder} level.\textsuperscript{137} In the UK, although self-regulation plays an important role, the Press Complaints Commission has been established.\textsuperscript{138} Even in the US, where free speech is interpreted most broadly,\textsuperscript{139} the law imposes some red lines for the traditional media: obscene broadcasts are banned outright, and indecent and profane broadcasts are only allowed under certain conditions. The rules on obscenity and indecency in the media are closely enforced by the Federal Communications Commission.\textsuperscript{140}

\textsuperscript{128} Ibid.
\textsuperscript{132} Kelly, Mazoleni and McQuail (eds.), \textit{The Media in Europe}, 49.
\textsuperscript{133} Ibid., 60.
\textsuperscript{134} Ibid., 26.
\textsuperscript{135} Ibid., 38.
\textsuperscript{136} Ibid., 133.
\textsuperscript{137} Ibid., 85.
\textsuperscript{139} For example, there’s no law in the US prohibiting Holocaust denial, the glorification of the Nazi Party or Communist terror. Such laws, by contrast, exist in almost all EU member states.
There’s a risk that social media platforms, faced with responsibility for the content they publish, might be tempted to overreact and start taking down lawful content on initial notification, without trying to verify the validity of the claim, simply in order to avoid potential liability. It is also entirely imaginable that anti-fake warnings could become a badge of honour for populist campaigns.\textsuperscript{141} This is already happening: in reaction to the Facebook flagging tool, some right-wing bloggers have spoken of the ‘pre-Thoughtcrime Unit [being] up and running’.\textsuperscript{142} To reduce this risk, content creators need to be guaranteed the right to appeal against the deletion. ‘The idea seems appealing; at least it merits a discussion. But there are several issues to consider. The Internet still is the symbol of freedom of speech, participative democracy and so on. Any attempt to regulate it excites an outcry’, noted Philippe Laloux.\textsuperscript{143}

There’s another theoretical angle to justifying the regulation of social media: one could consider them to be a twenty-first century public utility or public carrier. According to this theory, Facebook and Google are today about as optional as electricity, modern medicine or phones. They are essential to our social existence, almost as indispensable as the local sewage and water systems. Of course, there is a solid counter-argument to this. Companies such as Facebook and Google do not have a monopoly, not in the same way that the electricity grid does. In 2006 there was another brand that dominated the market—MySpace (now in oblivion). The history of the Internet is littered with fallen giants such as AOL, AltaVista and (the most recent example) Yahoo! The very nature of online technology guarantees that no social media monopolies can be created.\textsuperscript{144} But such a company can hold a dominant position, at least for a while.

\begin{flushright}
\textsuperscript{141} Stern, ‘Germany’s Plan to Fight Fake News’.
\textsuperscript{142} Hunt, ‘Disputed by Multiple Fact-Checkers’.
\textsuperscript{143} P. Laloux, interview.
\textsuperscript{144} A. Thierer, ‘The Danger of Making Facebook, LinkedIn, Google and Twitter Public Utilities’, \textit{Forbes}, 24 July 2011.
\end{flushright}
Where to regulate?

The national level seems to be the most appropriate for any social media legislation, for two reasons: first, press codes have traditionally been imposed by national governments and parliaments. Second, the rule of subsidiarity, enshrined in the EU treaties, implies that any problem should be addressed at the national level unless regulating it on the EU level brings added value. Given the differences between countries in terms of existing press laws and legal traditions, it would be next to impossible to create an EU-level, one-size-fits-all law.

So far, efforts undertaken to regulate social media on a global scale have failed. The 2012 UN World Conference of International Telecommunications formally adopted the ‘Final Acts’ (a 30-page-long set of proposed new regulations on telecommunication and Internet governance), but the legal nature of this document was non-binding. Of the 144 participating countries, only 89 signed the Final Acts. Many Western countries, including the US, Canada, France, the UK and Germany, refused to accept the document.\textsuperscript{145}

\textsuperscript{145} P. Wu, 'Impossible to Regulate', 284.
A long-term proposal: education
In the long term, improving Internet literacy might be the best and most democratic way to combat fake news and limit the impact of social media bubbles. There is a lot to be done. A study conducted in 2016 by Stanford University’s History Education Group suggests that the younger generation lacks critical-thinking skills and the ability to separate objective information from propaganda. For example, 80% of middle school pupils who took part in tests could not see the difference between a genuine news story and sponsored content, even if the latter was visibly labelled as ‘sponsored’. Many students, despite their presumed online fluency, were unaware of the basic indications of verified information.

Not only can e-literate citizens tell the truth from the fake; they also understand the consequences of their online activity and are aware that everything that they click on, like, share, pin, quote or retweet, in one way or another influences their audience. They understand that social media give them publishers’ powers by allowing them to decide what goes viral and what does not. Education would address the problem of fake news on the demand side: currently too many users forward content without reading anything but the headline, and too few care about the source of the content. This is a cultural problem, one that is shaped by disconnections in values, relationships and the social fabric. Several European governments seem to share this view. At the beginning of 2017, Czech Prime Minister Bohuslav Sobotka announced his government’s intention to modify the school curriculum in order to teach children how to assess the credibility of information. The Swedish government has announced a similar plan to teach children how to differentiate between reliable sources of information and unreliable ones. Social media platforms have also made efforts to educate their users on how to spot fake news. In April 2017 Facebook announced that on newsfeeds in selected countries (14 in total) it had published a special educational tool, a tutorial of sorts, created in cooperation with the non-profit organisation First Draft.

---

147 Borel, ‘Fact Checking’.
148 Boyd, ‘Google and Facebook’.
Conclusions and recommendations
Only a few years ago social media were being heralded as the hope for modern democracy: the ultimate tool to allow perfect public debate, where every voice counts and can be heard. To a large extent, this has proved true. As already noted, social media have become an essential part of our lives, offering numerous material and non-material benefits. Some scholars argue that the content published on social media platforms—including fake news!—is a true reflection of people’s emotions, and hence it is the essence of democracy. Instead of trying to regulate the platforms, they claim, mainstream politicians should make better efforts to convince people.\textsuperscript{152} Such views are well intentioned but they seem to be partially flawed. For example, should one consider any fake news story purposely created by an ill-intentioned foreign power to be a ‘true reflection of people’s emotions’? Should people unintentionally base their decisions on false news? Even \textit{The Economist}, a weekly opinion-forming journal known for its liberal stance, has recently concluded that the era of digital exceptionalism cannot last forever, as platforms’ influence over public life and the economy is becoming too great.\textsuperscript{153} The evidence at hand shows that the platforms’ very business models and construction unintentionally entrench echo chambers.\textsuperscript{154} Undeniably, in the longer term, education and improving the e-literacy of citizens are probably the best ways to limit the influence of lies. However, it will be many years until the effects of such policies kick in. In the meantime, Western democracies need quick fixes.

Bold steps are necessary. Fake news negatively influences the public debate and it must be stopped. The functioning of democracies is at stake. Fake news is as dangerous as hate speech and other illegal content.

Although voluntary measures to fight fake news, such as increasing the firepower of fact-checkers, should be encouraged, these are not sufficient. As explained, the existing press laws seem to offer the best available basis for a new approach. The risk of the excessive restriction of free speech must not be taken lightly.\textsuperscript{155} Social media must not be allowed to arbitrarily decide what is true and what is not. The deficits in the existing governance of Internet content are more than clear, but the solution cannot be to push even more decision-making about what can be said on the Internet onto unaccount-
able private corporations.\textsuperscript{156} This is exactly why there should be state-imposed regulation to guarantee that minimum standards are applied. For example, it should be ensured that fact-checking is done by humans, not by artificial intelligence. This might add substantial costs to the social media platforms’ business models. Regardless of how important it is that platform owners’ are comfortable with doing business—social media are a vital part of today’s economy, after all—the public interest carries much greater weight.

Therefore, this paper makes the following recommendations:

- Social media platforms should be given a limited amount of time, preferably one year, with a maximum of two years, to prove that they can effectively fight the spread of fake news on their own. The effectiveness of voluntary measures should be assessed by governments, political actors, non-governmental organisations, representatives of civil society and the media. If the assessment is negative (i.e. fake news is still spreading), states should impose top-down regulation.

- Governments must seriously consider applying the existing press laws to social media, with some necessary modifications due to the technology involved. A new procedure, tentatively called ‘notice and correct’, should be introduced in parallel to the existing ‘notice and take down’ procedure (which relates to illegal content posted online).

- The ‘notice and correct’ procedure should function in the following way:
  
  – Any interested party (physical or legal person) is entitled to notify the platform of content that it deems incorrect and harmful to its interests; once the party notifies the platform it should receive an obligatory confirmation of notification;

  – Once notified, the company (platform) is free to choose one of the following scenarios:
    
    1. to do nothing;
    
    2. to delete the item;

\textsuperscript{156} B. Wagner, interview.
3. to refer the item in question to fact-checkers (in-house or independent); based on their assessment, the platform decides whether to keep the item unchanged, to correct it itself, to ask the author (if identifiable) to correct it on penalty of deletion or to delete it.

If the company decides to delete or correct the item itself, the authors must be informed of the reasons for the intervention. Authors need to be allowed the right to appeal to an independent body (preferably the courts) should they consider that their rights have been violated.
Bibliography


Emmanouilidis, J. A. and Zuleeg, F., *EU@60—Countering a Regressive & Illiberal Europe*, European Policy Centre (Brussels, October 2016).


