
Disinformation

by Fazila Mat and Niccolò Caranti

Dossier ECPMF
January 2019

Index

"Fake news". An old term with new implications.....	3
Choosing the right term.....	4
Producing disinformation: a process encompassing several layers.....	7
IT Companies' Responsibilities.....	11
The reach of disinformation in Europe.....	12
Actions taken to tackle disinformation.....	17
The Response of the European Union.....	18
The Multi-stakeholder Process of the European Commission.....	19
Other Remedies to Disinformation.....	21

“Fake news”. An old term with new implications



Source: *The fin de siècle newspaper proprietor* by Frederick Burr Opper

The term “fake news” in itself is not new at all. In a [1894 illustration](#) by Frederick Burr Opper, a reporter is seen running to bring “fake news” to the desk.

“Fake news” and disinformation themselves are even older: a [short history](#) by the International Center for Journalists cites Octavian’s propaganda campaign against Mark Antony as an ancient example.

Overtime developments in technology, such as the invention of the printing press, enabled a faster, easier diffusion of information, including disinformation. In 1835, The Sun of New York published 6 articles about the discovery of life on the Moon, in what is now remembered as the “Great Moon Hoax”. In 1898, the USS Maine exploded for unclear causes, but American newspapers pointed to Spain, [contributing to the start of a Spanish-American War](#).

A [Council of Europe report](#) highlights that more recently, thanks to the Internet and social media, everyone can create and distribute content in real-time.

The term “fake news” [became ubiquitous](#) during the 2016 US presidential elections, being used by liberals against right-wing media and, notably, by the then candidate Donald Trump against critical news outlets such as CNN.

There are various [dangers](#) here. In 2016, a man opened fire in a restaurant in Washington, D.C., looking for a basement in which children were supposedly held prisoners: that was a fake news/conspiracy theory, known as [Pizzagate](#). Medical misinformation poses a threat to health; climate-related conspiracy theories pose a threat to the environment. Disinformation may continue to shape people’s attitudes even when debunked, and thus has [“real and negative effects on the public consumption of news”](#).

According to [a study](#) by David N. Rapp and Nikita A. Salovich, being exposed to inaccurate information can create confusion even when prior knowledge and experience should protect readers from considering and using it, and people show reliance on inaccurate information even after reading

fiction, or after conversations with people they have no particular reason to trust. More generally, “fake news” pollutes the information ecosystem: information is “as vital to the healthy functioning of communities as clean air, safe streets, good schools, and public health”. News is the raw material of good citizenship.

Distrust may be a by-product of ordinary misinformation, but some disinformation campaigns are explicitly aimed not at convincing someone of something, but at spreading uncertainty, sowing mistrust and confusion. As explained in the report [Lexicon of Lies](#) these campaigns are sometimes called “gaslighting”, a term originally used in psychology.

It is, however, important to note that mistrust is not only a consequence, but also a cause of disinformation: people turn to disinformation media because they do not trust mainstream media. This further increases distrust, and media distrust becomes a self-perpetuating phenomenon.

People who distrust the media are less likely to access accurate information. They will vote along partisan lines rather than consider the facts. The media are not able to perform their watchdog function, and as argued by the Digital, Culture, Media and Sport Committee of the British House of Commons in its [Interim Report](#), this poses a danger to democracy.

One must be aware that “fake news” accusations can also become a weapon in the hand of authoritarian regimes: a report by Article 19 underlines that world leaders use them to openly attack the media and, according to the Committee to Protect Journalists, 28 journalists imprisoned for their work worldwide have been charged with spreading false news: 11% of the 251 journalists detained globally.

Choosing the right term

“Fake news” has been variably used to refer to more or less every form of problematic, false, misleading, or partisan content.

A 2017 paper by Tandoc et al. examined 34 academic articles using the term “fake news” between 2003 and 2017, showing that it has been used to refer to six different things: satire, parody, fabricated news, manipulated or misappropriated images or videos, advertising materials in the guise of genuine news report, and propaganda.

The term “fake news” has thus been criticised for its lack of “definitional rigour”, and many chose not to use it, except between inverted commas. A Handbook by UNESCO even put a strikethrough on it in its cover.

According to Johan Farkas and Jannick Schou, “fake news” has become a “floating signifier” with at least three moments: a critique of digital capitalism, a critique of right-wing politics and media, and a critique of liberal and mainstream journalism. It is thus used as part of ideological battles to impose a specific viewpoint onto the world.

There have been many attempts to find alternatives to the term and systematize the conceptual framework.

Mark Verstraete, Derek E. Bambauer, and Jane R. Bambauer distinguish several types of “fake news” based on the motivation and on whether there is the intention to deceive readers (or not):

- satire: purposefully false content, financially motivated, not intended to deceive readers;
- hoax: purposefully false content, financially motivated, intended to deceive readers;
- propaganda: purposefully biased or false content, motivated by an attempt to promote a political cause or point of view, intended to deceive readers;
- trolling: biased or fake content, motivated by an attempt to get personal humor value (the “lulz”), intended to deceive readers.

Intent		
Deceive ²⁸ ← → Not Deceive		
Financial ↑ Payoff ↓ Not Financial	<i>Hoax</i> Example: Macedonian Teenagers	<i>Satire</i> Example: The Onion
	<i>Propaganda</i> Example: Controversy Re: Hillary Clinton's Health <i>Trolling (Lulz)</i> Example: Jenkem Episode	<i>Humor</i> Example: Twitter parody accounts ²⁹

Source: Identifying and Countering Fake News di Mark Verstraete, Derek E. Bambauer, and Jane R. Bambauer

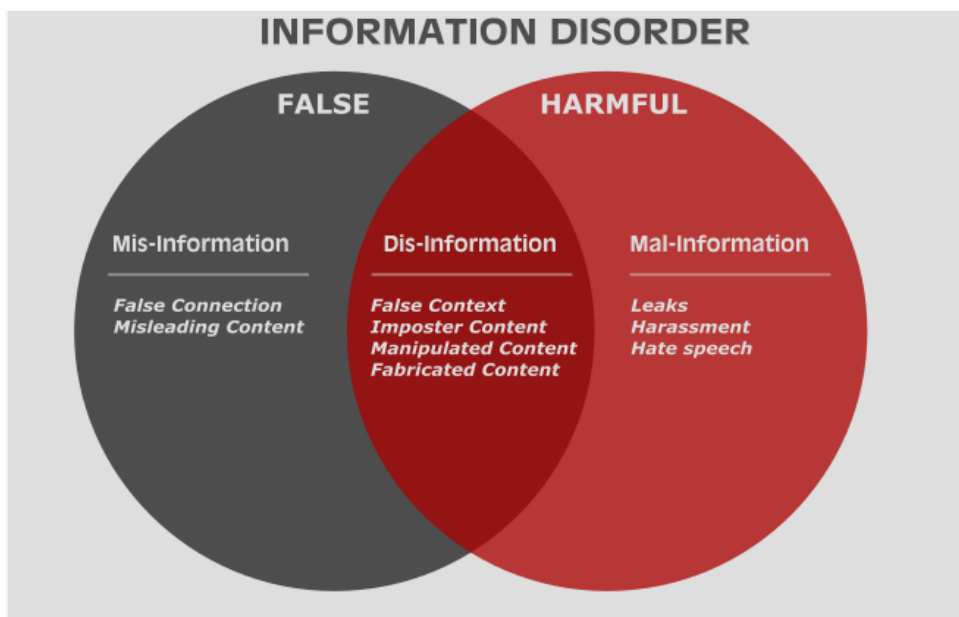
Claire Wardle, instead, [distinguishes](#) seven types of “fake news” on the basis of the motivation of the creators and the dissemination mechanisms: satire or parody, misleading content, imposter content, fabricated content, false connection, false context, and manipulated content.



Source: Fake news. It's complicated, by Claire Wardle, First Draft

Claire Wardle and Hossein Derakhshan in [their report for the Council of Europe](#) use the concept of “information disorder” and considering the motivation of the producer distinguish between:

- mis-information: false information shared without meaning any harm;
- dis-information: false information shared to cause harm;
- mal-information: genuine information shared to cause harm.



Re-elaborated graph from the original source: Claire Wardle and Hossein Derakhshan, *Information disorder: Toward an interdisciplinary framework for research and policy making*, CoE, 2017, p.25.

Producing disinformation: a process encompassing several layers

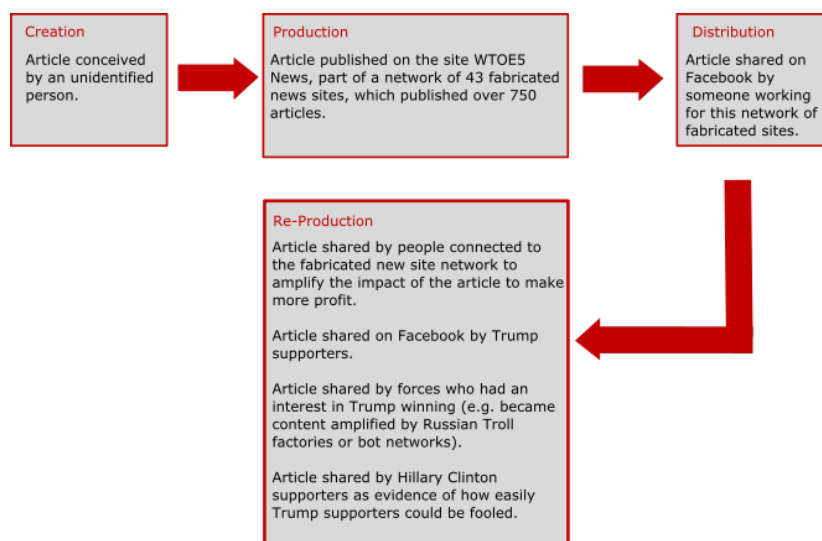
Disinformation can be produced in a variety of formats, including distortion of images and [deep-fakes](#), i.e. [manipulated audios and videos](#) that sound and look like a real person thanks to artificial intelligence (AI) techniques. As highlighted in [a recent report](#) written for the UK government, these examples will only become more complex and harder to spot as the software becomes more sophisticated.

Whatever the medium might be, Wardle and Derakhshan [identify](#) three elements (the agent; the message; the interpreter) and three phases (creation; production; distribution) of “information disorder” and stress the importance of considering the latter alongside the former.

The agents could be official actors (i.e. intelligence services, political parties, news organisations, PR firms or lobbying groups) or unofficial actors (groups of citizens that have become evangelised about an issue) who are politically or economically motivated. Social reasons - the desire to be connected with a certain group online or off - and psychological reasons can also play a role. The agent who produces the content is often fundamentally different from the agent who creates it and from the agents that distribute and reproduce it. As underlined by the two researchers, once a message has been created, it can be reproduced and distributed endlessly, by many different agents, all with different motivations. For example, a social media post shared by several communities could be picked up and reproduced by the mainstream media and further distributed to other communities.

• Virality

The emergence of new business models and the decrease of funds for quality media have paved the way for commercialising and sensationalising news, which is fertile ground for disinformation to flourish. Because shocking news draws greater attention, it also spreads more quickly on social media, thus becoming “viral”.



Using the example of the 'Pope Francis Shocks World, Endorses Donald Trump for President, Releases Statement' fabricated news articles to test the Three Phases of Information Disorder

Re-elaborated graph from the original source: Claire Wardle and Hossein Derakhshan, Information disorder: Toward an interdisciplinary framework for research and policy making , CoE, 2017, p.25.

According to a [research](#) conducted by Craig Silverman on the final three months of the US presidential campaign, the top “fake” election news stories generated more total engagement on Facebook than the top election stories from 19 major news outlets combined. The news about Pope Francis endorsing Donald Trump for the US presidency - first published by the WTOE5News.com site, which does not exist anymore - had the highest number of Facebook engagements (over 960,000 shares, reactions, and comments), whereas the top mainstream news story had only 849,000 engagements.

- **Clickbait**

As explained in the report [Why Does Junk News Spread so Quickly](#), the advertisement system provided by the self-service ad technology of companies such as Google and Facebook allows the creator of a webpage to profit for every click or impression per ad. Therefore, if a story drives e.g. on Facebook traffic to a site or is widely shared on social media, publishers that run the ads earn money each time the ad gets a click.

However, this business model could become particularly dangerous if false and harmful content is involved. [The case of Macedonian teenagers](#) who saw fake-news sites as a way to make money ahead of the 2016 US elections is an eloquent example. [BuzzFeed](#) identified over 100 pro-Trump websites run from the Macedonian town Veles and reported that the profits of these teenagers could reach “\$5,000 per month, or even \$3,000 per day”.

- **Trolling**

The Russiagate investigation aimed at verifying possible Russian influence in the 2016 US presidential election shed light on other elements and phases of information disorder. As [reported](#) by the New York Times, the Internet Research Agency (IRA) - also known as the “Russian troll factory” - made a weekly payment of \$1,400 to the recruits working as trolls for the agency. A troll is a real person who “intentionally initiates online conflict or offends other users to distract and sow divisions by posting inflammatory or off-topic posts in an online community or a social network. Their goal is to provoke others into an emotional response and derail discussions”, [reads](#) an analysis by the Digital Forensic Research Lab of the Atlantic Council.

Russian oligarch Yevgeny Prigozhin, who controlled two companies that financed IRA's operations - along with 12 other Russians and three Russian organisations - was charged by the US Justice Department for taking part in a wide-ranging effort to subvert the 2016 election and support the

Trump campaign. In June 2018, the United States House Intelligence Committee released a list of 3,841 Twitter usernames associated with the activities of Russia's "Troll factory". In a [recent working paper](#), D. L. Linvill and P. L. Warren analysed 1,875,029 tweets associated with 1,311 IRA usernames. The researchers identified a few categories of Twitter accounts, including Right Trolls (that spread nativist and right-leaning populist messages, often sending divisive messages about mainstream and moderate Republicans), Left Trolls (sending socially liberal messages and discussing gender, sexual, religious, and racial identity, attacking mainstream Democratic politicians) and Fearmongers (responsible for spreading a hoax about poisoned turkeys near the 2015 Thanksgiving holiday).

- **Users' cognitive and psychological impact**

The spread of these contents and rumors would not have been possible without the exploitation of psychological and cognitive mechanisms of social media users, who interpret the message. As explained in Craig Silverman's [Lies Damned Lies and Viral Content](#), people are frustrated by uncertainty and therefore tend to believe rather than question. In addition, if the rumor has personal relevance for us and we believe it to be true, we are more likely to spread it. Recent studies confirm that familiarity is a powerful persuasive factor, while repetition is one of the most effective techniques for getting people to accept manipulated content. An element that further enhances these mechanisms is self-confirmation, another mental shortcut people deploy when assessing the credibility of a source or message.

As pointed out by Wardle and Derakhshan in their report [Information disorder](#), social networks' functioning mechanism makes it difficult for people to judge the credibility of any message: firstly, because posts look nearly identical on the platforms; secondly, because social media are designed to focus on the story instead of the source; and thirdly, because people on social media tend to rely on their acquaintances' endorsements and social recommendations.

- **Bots**

These mechanisms, and particularly the repetition component, are exploited by bots. As highlighted in [#TrollTracker: Bots, Botnets, and Trolls](#), bots are automated social media accounts run by algorithms. They automatically react to manipulative posts, producing a false sense of popularity about content is created. This

can easily lead to create conformity among human agents who would then further distribute bots' messages. These can become even more widespread when influential people are tagged.

- **Filter Bubbles**

Another important element is the “filter bubble”. The term was coined by Internet activist Eli Pariser to refer a selective information acquisition by website algorithms, including search engines and social media posts. The expression is strictly related to the personalisation of such information through the identification of peoples’ “click” and “like” behaviour, location, search history etc. This mechanism leads to the creation of “bubbles” where Internet users only get the information that is in line with their profile, while other sources and information get filtered. Several studies have pointed out the dangers of such a selective information acquisition, warning that it enhances polarisation of society by creating echo chambers.

This is, for example, what emerged in Demos’ 2017 research [Talking to Ourselves? Political Debate Online and the Echo Chamber Effect](#). The study analysed Twitter data from 2,000 users who openly expressed their support for one of four political parties in the United Kingdom, finding similar patterns between supporters of different political parties. The report concluded that an echo chamber effect does exist on social media, and that its effect may become more pronounced the further a user sits from the mainstream. Another study investigated the echo chamber dynamics on Facebook. [The analysis](#) conducted by Walter Quattrociocchi and other researchers explored how Facebook users consume science news and conspiracy science news. By examining the posts of 1.2 million users, researchers found that the two news types “have similar consumption patterns and selective exposure to content is the primary driver of content diffusion and generates echo chambers, each with its own cascade dynamics”.

Other studies have come to different conclusions on filter bubbles and echo chambers. Recent [academic research](#) used a nationally representative survey of adult Internet users in the UK. The study found that users who are interested in politics and those with diverse media diets tend to avoid echo chambers. Another [study](#) by Richard Fletcher and Rasmus Kleis Nielsen examined incidental exposure to news on social media (Facebook, YouTube, Twitter) in four countries (Italy, Australia, United Kingdom, United States). Comparing the number of online news sources used by unintentional social media users (users who do not see it as a news platform) with people who do not

use social media at all, the research found that the incidentally exposed users use significantly more online news sources than non-users, especially if they are younger and have low interest in news.

A [survey](#) of 14,000 people in seven countries found that “people who are interested and involved in politics online are more likely to double-check questionable information they find on the Internet and social media, including by searching online for additional sources in ways that will pop filter bubbles and break out of echo chambers”. The [2017 Digital News Report](#), published by the Reuters Institute for the Study of Journalism points to a similar conclusion, stating that “echo chambers and filter bubbles are undoubtedly real for some, but we also find that on average users of social media, aggregators, and search engines experience more diversity than non-users”.

IT Companies' Responsibilities

A 2017 [report](#) by the Tow Center for Digital Journalism highlights how the advent of social media platforms and technology companies led to a rapid takeover of traditional publishers' roles by platforms and organizations including Facebook, Snapchat, Google, Twitter, and Whatsapp. These IT companies have evolved beyond their role as distribution channels and now control what audiences see and even the format and type of journalism that gets to thrive. On the other hand, "news companies are given up more of their traditional functions and publishing is no longer the core activity of certain journalism organisations", reads the same report.

There are large numbers of people accessing news and information worldwide through social media platforms and social messaging softwares. Facebook [declares](#) 1.49 billion daily and 2.27 billion monthly active users on average for September 2018. According to [Reuters](#) , 370 million of these users were in Europe in December 2017. The [Reuters Institute Digital News Report](#) 2018 interviewed a sample of 24,735 people in “selected markets” (UK, US, Germany, France, Spain, Italy, Ireland, Denmark, Finland, Japan, Australia, and Brazil): 65% of the respondents had used Facebook “for any purpose” in the week before the survey, while 36% of them had used it for news.

Facebook is governed by the News Feed algorithm, which has been modified several times to increase priority of videos, reduce priority of clickbait, emphasise family and friends, etc. Recently the platform altered its algorithm so that content identified as disinformation ranks lower. Facebook is way bigger

than Twitter and there is still loads of “fake news” there — and much more, in sheer numbers, than there is on Twitter, as punctualised by [NiemanLab](#) .

In 2014, Facebook launched Trending Topics, a section similar to Twitter trends. In 2016, [Gizmodo](#) quoted former Facebook employees saying that the section was curated to suppress conservative news. Months later, Facebook changed the way the section operated and it became fully algorithmically driven, until it was shut down in 2018 after accusations of spreading “fake news”.

According to a recent [working paper](#) by Stanford and New York University researchers, the platform’s attempts to get “fake news” and disinformation out of people’s feeds seem to be working. The problem, though, is that social media networks’ algorithms are not transparent: this means that all we know about it derives from statements by companies’ executives and attempts by third parties to reverse-engineer it, such as [the one by Fabio Chiusi and Claudio Agosti](#).

In November 2015, Google [introduced](#) a new policy, excluding AdSense from websites who misrepresent who their owners are and deceive people, including impersonating news organisations. Also other IT organisations are exploring ways in which content on the internet can be verified, kite-marked, and graded according to agreed definitions. Brands [may choose](#) to block advertisement on certain websites through blacklisting. Campaigns, like [Sleeping Giants](#) , push them to do that. On the other hand, the [report](#) of the independent [High level Group on Fake News and Online Disinformation](#) (HLEG) points out that many of these initiatives are only taken in a small number of countries, leaving millions of users elsewhere more exposed to disinformation. Furthermore, because of the scarcity of publicly available data, it is often hard for independent third parties (fact-checkers, news media, academics, and others) to evaluate the efficiency of these responses.

The reach of disinformation in Europe

Most existing data and studies about “fake news” refer to the United States, especially to the 2016 presidential elections. Nevertheless, there have been concerns in several other countries, including European ones, particularly around the 2016 Brexit referendum in the UK and other European electoral appointments.

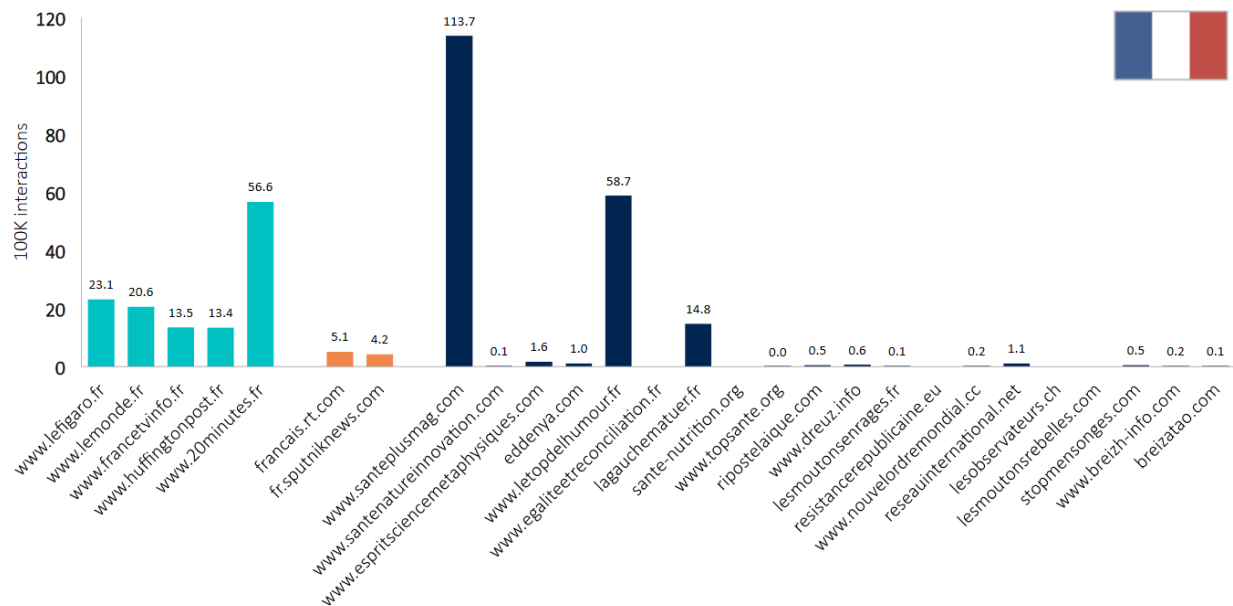
In the United Kingdom, the Digital, Culture, Media and Sport Committee of the House of Commons is carrying out an inquiry on disinformation and “fake news”. An [interim report](#) published in July 2018 discusses the role of Russia in supporting the campaign for Brexit, planting “fake news” in an attempt to “weaponise information” and sow discord in the West.

Academics Chadwick, Vaccari and O’Loughlin [analyzed](#) the results of a survey conducted with 1,313 individuals in the United Kingdom, highlighting that during the 2017 UK general election campaign 67.7% of the respondents admitted to having shared news that were exaggerated or fabricated. Moreover, they found that there is a positive and significant correlation between sharing tabloid news articles and exaggerated or fabricated news: authors estimate that there is a 72% probability of dysfunctional news sharing by respondents sharing one tabloid news story per day. The study also shows that “the more users engage with politically like-minded others online, the less likely it is that they will be challenged for dysfunctional behavior”, with someone disputing the facts of the article or claiming it is exaggerated.

A [fact sheet](#) by the Reuters Institute for the Study of Journalism based on research in France and Italy shows that so-called “fake news” has a limited web reach: most reaches less than 1% of the online population in both countries, in comparison Le Figaro in France had an average monthly reach of 22.3%, La Repubblica 50.9%. Moreover, far less time is spent on false news websites than on mainstream news websites: people spent an average of 178 million minutes per month with Le Monde, and 443 million minutes with La Repubblica, while the most popular false news websites in France were viewed for around 10 million minutes per month, and for 7.5 million minutes in Italy.

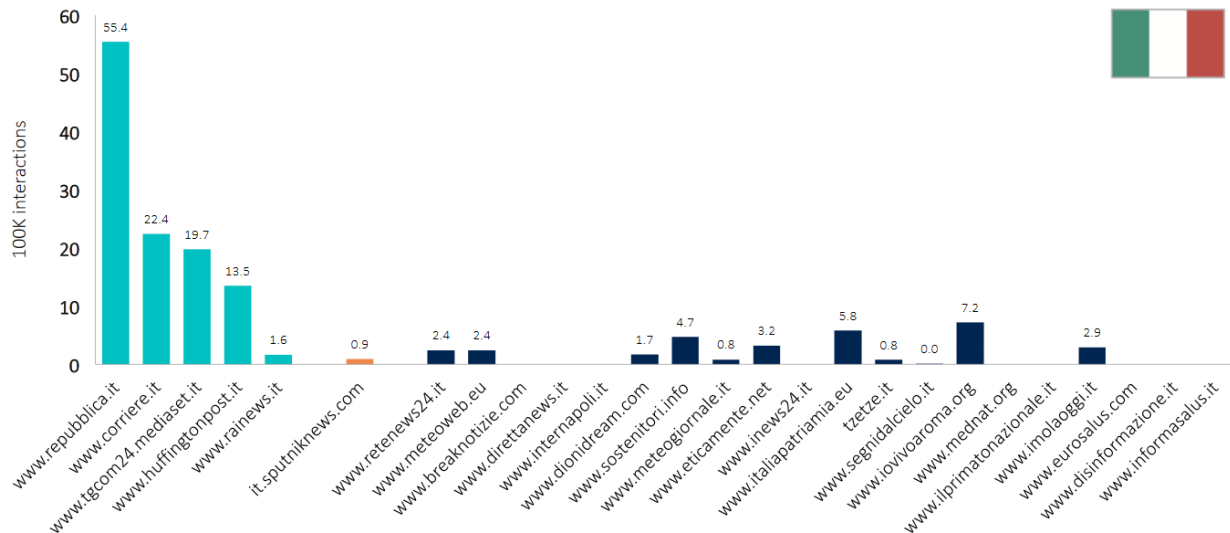
However, the fact sheet also shows that in France “a handful of false news outlets (...) generated more or as many interactions as established news brands”, but they are the exception, as most of them generated less interactions. In Italy, there are 8 false news websites that generated more interactions than the Rainews website - which is however not widely used - and they are far behind La Repubblica and Il Corriere della Sera.

Figure 3. Average monthly Facebook interactions for prominent French news sites, and some of the most popular false news sites (2017)



Source: Measuring the reach of "fake news" and online disinformation in Europe by Richard Fletcher, Alessio Cornia, Lucas Graves, and Rasmus Kleis Nielsen

Figure 6. Average monthly Facebook interactions for prominent Italian news sites, and some of the most popular false news sites (2017)



Source: Measuring the reach of "fake news" and online disinformation in Europe, by Richard Fletcher, Alessio Cornia, Lucas Graves and Rasmus Kleis Nielsen

Other studies suggest similar results. Gaumont, Panahi and Chavalarias [reconstructed](#) the political landscape of France on Twitter during the 2017 presidential election: authors compared the links that were shared on Twitter with those that were classified as "fake news" by Le Monde and found that "fake news" is not heavily shared in France by people interested in politics. According to a European Union Institute for Security Studies' [paper](#) on Russian cyber strategies, the Macron leaks (the release of gigabytes of data hacked from Emmanuel Macron's campaign team two days before the final round of the presidential election) were not able to reach mainstream sources of information. According to Jeangène Vilmer, this was because of various reasons, including structural ones (the two-round direct elections of the president, a resilient media environment, and cartesianism) and mistakes by hackers (overconfidence, timing, and cultural clumsiness).

Russia [has been accused](#) of using "fake news" as information warfare abroad, as analyzed in the monograph by Popescu and Secieru published by the European Union Institute for Security Studies. A recent [study](#) on "information manipulation" by the French government states that the 80% of the European authorities consulted attribute influence efforts in Europe to Russia. Other source of influence: "comes from other States (mainly China and Iran) and non-state actors (Jihadist groups, in particular ISIS)", the report reads. Some scholars, however, have questioned this percentage on the basis that such a volume is hardly measurable because there is no consensus on what exactly constitutes disinformation.

In Italy, EU DisinfoLab has [tested](#) a disinformation detection system on Twitter in the run-up to the March 2018 elections, finding only a few examples of disinformation and no evidence of meddling by foreign actors. Besides, AGCOM, the Italian telecommunications regulatory body, launched an [inquiry](#) about "online platforms and the information system". The study found that there was a peak in the quantity of disinformation during the electoral campaign: while throughout April 2016 fake content was an average of 1% of the total (2% if considering only online content), it reached an average of 6% in the 12 following months (10% if considering only online content). It also found that "fake news" has a shorter life cycle than real news: on average, real news lasts 30 days, while fake news' life is limited to 6 days (duration is calculated as the average distance between the first and the last day in which a piece of news appears at least once).

A group of researchers from Oxford University [studied](#) Twitter data on bot activity and junk news using a set of hashtags related to the 2017 German Parliamentary Election. The scholars took into examination a dataset containing approximately 984,713 tweets generated by 149,573 unique users, collected between the 1st and 10th of September 2017, using hashtags associated with the primary political parties in Germany, the main candidates, and the election itself. The researchers found that the traffic generated by the far-right Alternative für Deutschland (AfD) accounts was much higher than party's share of voter support. They also found that most political bots were working in the service of the far-right AfD, although the political bots' overall impact was minor. The study also highlighted that German social media users shared four links to professional news sources for every link to junk news. Figures are much higher when compared with users' sharing habits in the US, but lower than in France.

Another [report](#) published in 2018 by the Knight Foundation confirms such data. Bradshaw and Howard calculated the ratio of professionally-produced content to junk news (which includes hate speech, hyper partisan content, etc.) by number of links in the United States, Germany, France, and the United Kingdom: they found a 1:1 ratio in the US, while the situation was better in the three European countries.

Country	Details	Ratio of Professionally-Produced Content to Junk News, by Number of Links
USA	Sample taken between November 1-11, 2016, totaling 1,275,430 tweets.	1:1
Germany	Sample taken before voting, from September 1-10, 2017, totaling 149,573 tweets.	4.4:1
France	Sample taken before voting, from April 27-29, 2017, totaling 76,063 tweets.	6.5:1
United Kingdom	Sample taken before voting, May 27-June 2, 2017, totaling 27,059 tweets.	4.7:1

Source: Authors' calculations, based on results from: Desigaud et al 2017, Gallacher et al. 2017, Howard et al. 2017c, and Neudert et al. 2017.

Source: *Why Does Junk News Spread So Quickly Across Social Media?* Di Samantha Bradshaw e Phillip N. Howard

In Ireland, the government established an Interdepartmental Group on the Security of Ireland's Electoral Process and Disinformation, which published its [first report](#) in July 2018. According to a survey it conducted, 57% of the respondents are concerned about “fake news”, but only 28% understand the role of algorithmic targeting and spread of disinformation.

A [study](#) carried by the [Open Society Institute of Sofia](#) shows that there is a clearly geographic pattern in the potential of resilience “fake news” and associated “post-truth” phenomenon: the countries with a better performance are in the [North and Northwest of Europe](#), as opposed to the countries in the Southeastern Europe, while countries like Hungary, Italy, and Greece are in a middle cluster.

Actions taken to tackle disinformation

The unprecedented proliferation of fraudulent information and its potential to spread without limits though the digital space have led several international and European organisations to develop strategies to tackle the phenomenon. However, the task is particularly difficult for a number of reasons. First of all, as we have illustrated before, the phenomenon includes a range of different types of disinformation and, without an agreement on the phenomenon to manage, developing a common strategy is difficult. Furthermore, the attempts to manage it could pose serious risks of curbing freedom of expression. In this regard, identifying the authority in charge of spotting disinformation emerges as one of the most insidious tasks. Other challenging questions concern whether such power should be in the hand of private technology companies such as Facebook and Twitter or a responsibility of state authorities. And finally, in the latter case, whether this authority should be the government or the judicial authority. On the other hand calls for directing efforts in order to enhance quality information and citizens’ media and information literacy have also become central to the issue.

In the [Joint declaration on freedom of expression and “fake news”, disinformation and propaganda](#) adopted on March 3rd, 2017, international bodies concerned with freedom of expression - including the UN Special Rapporteur on Freedom of Opinion and Expression and the OSCE Representative on Freedom of the Media - addressed these questions inviting all relevant actors - states, intermediaries, media, journalists, and civil society – to adopt tailored measures to counteract the spread of “fake news”. The declaration addressed in particular the role of states in establishing a clear regulatory framework that protects the media against political and commercial interference; in

“ensuring the presence of strong, independent and adequately resourced public service media”, as well as in providing “subsidies or other forms of financial or technical support for the production of diverse, quality media content”. It was also stressed that fight against “fake news” cannot result in unnecessary limitations of freedom of expression. Therefore, states should limit technical controls over digital technologies such as blocking, filtering, and closing down digital spaces as well as any efforts to “privatise” control measures by pressuring intermediaries to take action to restrict content. The document highlighted the need for transparency and rules prohibiting undue concentration of media ownership and the role of media literacy as a key tool to strengthen the public’s ability to discern, to be actively supported by public authorities.

The Response of the European Union

The European Union’s first actions to contrast the “fake news” phenomenon have been specifically intended to address Russia’s disinformation campaigns. Following the launch of the European External Action Service [East StratCom Task Force](#) , in March 2015, the European Council [tasked](#) the High Representative in cooperation with EU institutions and Member States with submitting an action plan on strategic communication, stressing the need to establish a communication team as a first step. In May 2016, the European Commission (EC) presented its [communication on online platforms](#) to the EP, the Council, the European Economic and Social Committee, and the Committee of the Regions. On 12 September 2017, the Task Force presented its new EU vs Disinfo [website](#) , described as part of the “EU vs Disinformation” campaign “to better forecast, address and respond to pro-Kremlin disinformation”.

A recent [analysis](#) by Alemanno inquired into the functioning mechanisms of [Disinformation Review](#) , defined as “the flagship product of the EU vs Disinformation campaign” in the EU vs Disinfo website. Alemanno submitted a request for documents to the EEAS to understand how it works the reporting of disinformation articles to EU officials by a network of over 400 experts, journalists, officials, NGOs, and think tanks in more than 30 countries. Alemanno concluded that the review process of the sources of information placed on the Disinformation Review and the criteria used to debunking “fake news” and Russian propaganda is vague and subjective. In March 2018, Alemanno et al. presented a complaint to the European Ombudsman highlighting the “ad hoc” fact-checking methodology of the

Disinformation Review that does not correspond to the one adopted by the international fact-checking community, led by the International Fact-Checking Network (IFCN). In addition, it is stressed the publication violated the rights to freedom of expression and due process of those accused of distributing disinformation.

On the same month, the Dutch parliament [called](#) for the website's closure because it had wrongly listed articles published by Dutch media in its collection of cases conveying a “partial, distorted or false view or interpretation and/or spreading key pro-Kremlin messaging”. The taskforce removed the articles and the case was withdrawn. The debate continued with sixteen international affair analysts taking the side of the European service in an [op-ed](#) published on EUobserver. The op-ed claimed that the erroneous listing of the Dutch articles was the result of financial and manpower lack of EUvsDisinfo and stressed that “in view of the serious threat of Russian disinformation to Western democracies” the EUvsDisinfo site needed to be strengthened, not closed.

The Multi-stakeholder Process of the European Commission

The European Commission (EC) launched in November 2017 a multi-stakeholder process “to find the right solutions consistent with fundamental principles and applicable coherently across the EU” in view of the European elections of May 2019. The process started with a [two-day multi-stakeholder conference on fake news](#) and was followed by the launch of a [public consultation](#) that received 2,986 replies, with the largest number of reactions coming from Belgium, France, the United Kingdom, Italy, and Spain and high participation in Lithuania, Slovakia, and Romania. The first meeting of the High Level Expert Group on Fake News and Disinformation (HLEG) was also held on the same month. As a result of three further meetings in 2018, the HLEG produced a [report](#) in March 2018.

The HLEG made several recommendations, including the introduction of a code of principles for online platforms and social networks, ensuring transparency by explaining how algorithms select news, as well as improving the visibility of reliable, trustworthy news and facilitating users' access to it, as well as support for quality journalism across Member States to foster a pluralistic media environment and improving media literacy through support for educational initiatives and targeted awareness campaigns. The HLEG's report was followed in April 2018 by the [Communication](#) of the EC, which turned the 10 principles of the HLEG's report into 9 clear objectives “which should guide

actions to raise public awareness about disinformation and tackle the phenomenon effectively, as well as the specific measures which the Commission intends to take in this regard". The objectives range from improving the scrutiny of advertisement placements with the specific intent of restricting targeting options for political advertising to ensuring transparency about sponsored content; facilitating users' assessment of content through indicators of the trustworthiness of sources; improving the findability of trustworthy content; establishing clear marking systems and rules for bots; empowering users with tools so as to facilitate diversification on news sources; providing users with easily-accessible tools to report disinformation; ensuring that online services include, by design, safeguards against disinformation and finally providing trusted fact-checking organisations and academia with access to platform data, while respecting user privacy, trade secrets, and intellectual property.

In September 2018, a self-regulatory [Code of Practice](#) to address the spread of online disinformation and "fake news" was agreed upon by representatives of online platforms, leading social networks, advertisers, and the advertising industry. The Code also includes an [annex](#) identifying best practices that signatories will apply to implement the Code's commitments. According to the new code, among other things, the advertising revenue of businesses spreading "fake news" should be banned and the fight against fake accounts and bots should be intensified. Several IT companies including Google, Facebook, Twitter, and Mozilla stated their willingness to comply with self-regulation as proposed by the EU, presenting individual [road-maps](#) to achieve the goal. On the other hand, the EU would monitor their effectiveness on a regular basis.

However, the "sounding board" - the Multistakeholder Forum on Disinformation Online group comprising ten representatives of the media, civil society, fact checkers, and academia - have been [highly critical](#) about this Code. The board highlighted that the Platforms lack of a common approach as well as clear and meaningful commitments, measurable objectives or key performance indicators (KPIs). It has been argued that, given the impossibility to monitor the process, the code cannot be a compliance or enforcement tool and "it is by no means self-regulation". In December 2018, the EU presented an [Action Plan](#) to step up efforts to counter disinformation in the lead up to the continent-wide vote in the spring, with Andrus Ansip, Vice-President responsible for the Digital Single Market, singling out Russia as a "primary source" of "attempts to interfere in elections and referenda". The Action Plan focuses on four areas: improved detection; coordinated

response; online platforms industry; raising awareness and empowering citizens. The measures include beefing up the EEAS' strategic communication budget from €1.9 million in 2018 to €5 million in 2019, adding expert staff and data analysis tools, a new "rapid alert system" to be set up among the EU institutions and Member States to facilitate the sharing of data and assessments on disinformation. The EU authorities also want IT companies to submit reports from January until May on their progress in eradicating disinformation campaigns from their platforms. The companies are further expected to provide updates on their cooperation with fact-checkers and academic researchers to uncover disinformation campaigns.

Other Remedies to Disinformation

A group of four experts, academics and fact checkers, all members of the HLEG on Fake News, [have published a comment](#) on the EC's report on disinformation. They stressed that while disinformation is clearly a problem, its scale and impact, associated agents and infrastructures of amplification need to be further examined. There is a general understanding among scholars and experts that, unless clear proposals without significant downsides appear, disinformation should be treated with a "soft power approach" by the authorities.

Nonetheless over the last years there have been various ways to address the issue at the Member States level, including different regulatory approaches. One of the first regulatory actions against illegal online content was taken by Germany through the [Network Enforcement Law \(NetzDG\)](#) , which entered into force on January 1st, 2018. The law requires social platforms to remove content considered illegal within 24 hours and prescribes sanctions up to 50 million Euros for social networks that fail to act accordingly. The measure has been [criticised](#) because it triggers "the risk of greater censorship", as David Kaye, UN's Special Rapporteur on Freedom of Expression [put it](#) .

In February 2018, French president Emmanuel Macron [announced](#) his plan to [introduce legislation](#) that would curb the spread of disinformation during the country's future election campaigns. The president said that this goal would be achieved by enforcing more media transparency and blocking offending sites. He cited "thousands of propaganda accounts on social networks" spreading "all over the world, in all languages, lies invented to tarnish political officials,

personalities, public figures, journalists” and called for “strong legislation” to “protect liberal democracies”.

An additional control mechanism during the 2018 general elections was [introduced in Italy](#), where previous attempts of introducing a law against “fake news” were shelved after being criticised for being a censorship mechanism. Before the 2018 elections, the Italian State Police launched on its website a “red button” to signal fake news, on the initiative of then Minister of the Interior Marco Minniti. The protocol adopted on the eve of political elections, enabling the Postal Police to fact-check and report contents, has caused [concern](#) in the public debate.

In his [Taxonomy of anti-fake news approaches](#) , Alemanno argues that proposed anti-fake news laws “focus on the trees rather than the forest”, adding that “as such, they will not only remain irrelevant but also aggravate the root causes of the fake news phenomenon”. The solution prospected in the article is to “swamp fake news with the truth”. In explaining this idea, the scholar suggests the possibility of introducing a law that would require all social networks to invite readers to have easier access to additional perspectives and information, including articles by third-party fact checkers. This would only systematise a practice already implemented by Facebook on a voluntary basis with its [“Related Articles” feature](#) , argues the expert.

Human and computerised fact-checking has indeed become one of the main methods employed to contrast the spread of disinformation. The term is usually referred to internal verification processes that journalists put their own work through, but fact-checkers (or debunkers) dealing with disinformation are involved in ex-post fact-checking. The International Fact-Checking Network at Poynter has elaborated a [code of principles](#) signed by the majority of major fact-checkers which includes five commitments: Nonpartisanship and Fairness, Transparency of Sources, Transparency of Funding & Organization, Transparency of Methodology, and Open & Honest Corrections Policy. However, as explained by [Silverman](#), facts alone are not not enough to combat disinformation. Particularly challenging for fact-checking is the backfire effect - not only because contradicting people’s beliefs tends to cause them to double down on existing beliefs, but also because denying a claim helps it take hold in a person’s mind. Commenting on the European Commission’s call for the creation of an independent European Network of fact-checkers, the joint report [Informing the ‘Disinformation’ Debate](#) by EDRI raised questions based on the fact that “assuring the independence

and criteria for correctly carrying out 'fact-checking' is not as easy or simple as it sounds. There are risks of conflicts of interest, both direct and indirect, abuse of power, bias and other significant costs".

Many other solutions to remedy the propagation of disinformation have been proposed, from initiatives for publishers to signal their credibility, to technologies for automatically labelling misinformation and [defining and annotating credibility indicators in news articles](#). However, proposals like "credibility indexes" have also been criticised as methods weakening freedom of expression. In [Oxygen of Amplification](#), Phillips argues that in many cases amplification is indeed the objective of the action being reported. The scholar argues that this does not mean that the rules of newsworthiness should be sacrificed to avoid amplification, but when dealing with objectively false information, manipulation, and harassment campaigns some attention should be paid, as it is already done with coverage of suicides.

The role that traditional media can have in countering disinformation has been highlighted in several studies. As highlighted in [Lies, Damn Lies and Viral Contents](#) one of the reasons why "fake news" can reach a wider audience and "become viral" is due to many legitimate news websites' practice to publish unverified rumors and claims (sometimes coming from other media outlets) without adding reporting or value, thus becoming easy marks for hoaxsters and the like. If trust in traditional media has declined in recent years, this has to do with this phenomena.

European Federation of Journalists' general secretary Ricardo Gutiérrez, also a member of the HLEG on Fake News, [argues](#) with many others that "the best antidote to disinformation is quality journalism, [ethical commitment](#) and critical media literacy". It is widely acknowledged that a key mitigation factor for disinformation is media literacy. The HLEG Report stressed the importance of media and information literacy, because it enables people to identify disinformation. It also states that "media literacy cannot [...] be limited to young people but needs to encompass adults as well as teachers and media professionals" in order to help them keep pace with digital technologies.

[In another report](#) by Lessenski, media literacy is referred to "as a means to gauge the potential for resilience to the negative effects of diminishing public trust, polarisation in politics and society, and media fragmentation". And again, education and media literacy are key to building resilience to the post-truth phenomena and to reduce polarisation, while enhancing trust in society and in the media.

This publication has been produced within the project European Centre for Press and Media Freedom, co-funded by the European Commission.

Osservatorio Balcani e Caucaso Transeuropa

Operational unit of the International Cooperation Centre

www.balcanicaucaso.org

redazione@balcanicaucaso.org

