STUDY

Requested by the IMCO Committee

European Parliament

Societal costs of "fake news" in the Digital Single Market

Policy Department for Economic, Scientific and Quality of Life Policies
Directorate-General for Internal Policies
Author: Prof. Divina Frau-Meigs
PE 626.087- December 2018

EN

Societal costs of "fake news" in the Digital Single Market

Abstract

This study explores the mechanisms of "fake news" and their societal costs in the Digital Single Market. It describes the risks to the integrity of information and to the integrity of elections. It highlights the roles of the various actors involved in the production and amplification of such information disorders. Finally, it outlines responses that are being tested in different parts of Europe to deal with the issue.

The document has been provided by Policy Department A at the request of the European Parliament Committee on the Internal Market and Consumer Protection.

This document was commissioned by the European Parliament's Committee on the Internal Market and Consumer Protection.

AUTHOR
Prof. Dr Divina FRAU-MEIGS, University Sorbonne Nouvelle, Paris, France

ADMINISTRATOR RESPONSIBLE
Mariusz MACIEJEWSKI

EDITORIAL ASSISTANT
Irene VERNACOTOLA

LINGUISTIC VERSIONS
Original: EN

ABOUT THE EDITOR
Policy departments provide in-house and external expertise to support EP committees and other parliamentary bodies in shaping legislation and exercising democratic scrutiny over EU internal policies.

To contact the Policy Department or to subscribe for updates, please write to:
Policy Department for Economic, Scientific and Quality of Life Policies
European Parliament
B-1047 Brussels
Email: Poldep-Economy-Science@ep.europa.eu

Manuscript completed: November 2018
Date of publication: January 2019

This document is available on the internet at:
http://www.europarl.europa.eu/supporting-analyses

For citation purposes, the study should be referenced as: Frau-Meigs, D., *Societal costs of "fake news" in the Digital Single Market*, Study for the Committee on the Internal Market and Consumer Protection, Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament, Luxembourg, 2018.

CONTENTS

LIST OF ABBREVIATIONS

AMP	Accelerated Mobile Pages
API	Application Programming Interface
CBS	Corporate Broadcasting Service
CPM	Cost per Mile
CSA	Conseil Supérieur de l'Audiovisuel
DDoS	Distributed Denial Of Service
GAFAM	Google, Apple, Facebook, Amazon, Microsoft
GDPR	General Data Protection Regulation
H2020	Horizon 2020 work programme
IFCN	International Fact-Checking Network
MAS	Media Accountability System
MCN	Multi Channel Network
MIL	Media and Information Literacy
NATO	North Atlantic Treaty Organisation
OSCE	Organization for Security and Cooperation in Europe
PAC	Political Action Committee
PISA	Programme for International Student Assessment
PR	Public Relations

LIST OF BOXES

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

The "fake news" phenomenon has come to the centre of public and political attention since 2016, with the suspicion that several electoral events (elections in the U.S. and in France as well as referenda in the UK and Catalonia) have been tampered with. In all cases, the role of social media platforms as major vectors of such disruptive phenomena has been questioned, especially with scandals such as Cambridge Analytica's use of data from Facebook, and court cases such as Internet Research Agency vs. United States of America. As new elections are approaching in European countries, in particular the European Parliament´s elections in May 2019, it is urgent and necessary to evaluate the extent of the situation and adjust the necessary responses, especially in relation to trust in the Digital Single Market.

The purpose of this in-depth analysis, provided at the request of the European Parliament's Committee on the Internal Market and Consumer Protection, is to set the general framework for understanding the issue of "fake news" and how it can affect trust in the Digital Single Market. The paper offers an explanation of the phenomenon of "fake news" and indicates its evolution, from past processes to new specific incidents, differentiated from neighbouring concepts, such as propaganda or conspiracy. It provides for some elements of categorisation of "fake news" and illustrates their specificity in the XXIst century both with regard to the societal weaknesses that make Europe especially vulnerable to "fake news" and communication innovations that drive them. It redefines "fake news" as "malinformation", i.e. the triple conjunction of human malevolence, marketing malpractice and engineered malware. It situates malinformation in the family of information disorders such as misinformation, online radicalization and hate speech.

The analysis also reflects on the societal costs of this malevolent combination. Among these, there are economic costs but also, most important here, two major societal costs: the integrity of information and the integrity of elections.

As regards integrity of information, the paper points to the numerous asymmetries that exist not only between mainstream mass media and social media platforms but also between social media in the U.S. and in Europe. The European digital market has not been able to create search engines or social networks large enough to compete with the American ones. Many asymmetries between American and European companies are immediately apparent: tax obligations, social responsibilities, advertising constraints and public service obligations (like diversity, pluralism, and protection of minors). So nothing hinders the American-based social media from spreading malinformation to increase traffic and thus profits. These asymmetries increase risks of harm to legitimate information as it becomes extremely costly to produce and puts quality reference media under pressure to yield more verification of news, which in turn increases costs and adds time delays. All this takes place in a context of increasingly concentrated ownership and pressure for profits from share-holders and directors who consider information to be an appendage to their larger sectors of activity which can include construction, defence or other economic sectors without long-standing interest or competence in news.

As regards the integrity of elections, the paper describes the emergence of "hybrid threats" and the increased capacity of malign foreign intervention into EU affairs as the tension grows between democratic regimes and authoritarian regimes. Various phases of malinformation are described as well as the e-repertoire of strategies that recombine different mechanisms of malinformation into hybrid threats. Among the most characteristic elements are the drip dissemination of topics likely to create viral trends and polarization of topics well before the election time; the hack and delivery of strategic data to cause surprise or panic at determined stages of a campaign; the whitewashing of "fake news" via social media to promote amplification via mass media and conversely, and opportunistically, secret

collusion or objective alignment of interests with extremist or minority groups in a given country allowing internal and external actors to buttress each other.

"Fake news" thus creates challenges for both the sovereignty of Member States and for the cohesion of the European Union that are complex and may have far-reaching consequences. "Fake news" leads to information disorders related to the "transverse effect" of malinformation, i.e. the trans-border, trans-media, trans-lingual and trans-time nature of Internet networks. The direct and indirect impact of "fake news", though not proven yet by various nascent research protocols, may lead to long-term counter veiling deterrence measures which will reduce freedom of speech and the freedom to receive and impart information. Thus efforts to inhibit "fake news" will damage democratic processes. The overall democratic cost, hard to prove because it is dissipated and often indirect, is lack of trust in the reliability of institutions and media, with the added paradox that social media are the least trusted while being the most used.

The paper indicates how such costs can be minimised by identifying potential negative consequences of insufficiencies, areas for improvement and steps that need to be taken for the most efficient course of policy action in the area of information disorders at large. Recommendations are formulated on the basis of the analysis of existing responses. Measures are proposed for two stages of amplified malinformation flow: at the source before the event, when "fake news" is produced and later, when they are disseminated for consumption. The types of actors and of responses are not necessarily related. At the source, self-regulation is the preferred action for the private sector of mass media and social media alike while regulation is the favoured solution for the public sector at the moment of transmission for consumption, at the mouth, so to speak, of this flow of information. Education is the response most wanted by the civic sector and the users. Three high priority measures can be distinguished to combat "fake news", that do not recommend "hard" legislation (which might lead to censorship) but rather "soft" and "smart" action (guidelines, codes of good conduct, good practices, etc.), together with accountability, proportionality and revision mechanisms:

- Self-regulation: fact-checking is the new response of the mass and social media.
- Regulation: transparency is the main request of the public sector.
- Education: media and information literacy is fostered to ensure resilience and preparedness among the population at large as well as decision-makers working for political candidates and political parties.

The paper concludes by considering the current "trap-situation" created by malinformation as a harbinger of a new phase in co-regulation and re-intermediation of the Internet. Reactions to this "trap" will affect the European Digital Single Market whose overseers will need to coordinate actions and readjust directives and treaties in order to cope with this ever more urgent phenomenon. The EU should allocate specific means to deal with the responses suggested and focus on the transborder, pan-European scope of such mechanisms while paying attention to inter-sectional arrangements. It recalls that data and algorithms are not devoid of human prejudices, biases and should be monitored publicly to avoid unwanted and unexpected social disruptions.

Action at the Pan-European level (EU):

- Raise awareness about the multi-faceted dimensions of "fake news", "malinformation" and other "information disorders";
- Develop scenarios for the future, based on the concept of "information disorders" and promote cooperation with all actors and sectors;
- Reduce the asymmetries in regulation and public obligations between mass media and social media;
- Identify the risk of harm to information and develop strategies to stifle it at the source and deflect it at the moment diffusion, or mouth;
- Provide the basis for a core media and information literacy package, with requirements for skills, competences and evaluations, to be used across Europe, via PISA mechanism;
- Develop the general debate around the public interest of the Internet and social media and elicit interest and research in fact-checking strategies;
- Fund more research specifically targeted at information disorders, with international and interdisciplinary frameworks.
- Popularize cyber-security techniques and training, to ensure that decision-makers, but also citizens at large, are aware of strategies related to data, automation and computation (hacking, fraud…);
- Create a Europe-wide research institute on information disorders and promote digital citizenship education through its activities and tools.

Action for the mass and social media:

- Promote the social responsibility of media organisms to extend and adapt their self-regulatory systems (codes of good practice, information ombudsmen, press councils, fact-checking principles…) in relation to the Internet;
- Foster transparency at all levels, especially in terms of advertising revenue and its sources, algorithms, ranking and filtering design, particularly during political campaign periods;
- Provide researchers with easy access to data from social media and facilitate their cooperation with design engineers;
- Elaborate tools for fact-checking that can be extended to the whole community, with plug-ins for schools and libraries.

Action for key non-state actors:

- Map-out the professions at the interface between users and contents (like ombudsmen, webmasters, list moderators, librarians…) and train them in media and information literacy, including malinformation and information disorders;
- Encourage the social responsibility of the private sector, especially Internet service providers, operators and platforms;
- Train policy-makers at all levels of the decision-making process to sensitize them and arm them to fight "fake news" in their campaigns and everyday life so as to foster trust.

1. INTRODUCTION

Dissecting "fake news" is an interesting exercise in disruption analysis. From Pizzagate to Macron-Leaks to fake-vax, it builds on collective "liquid fears" as Zigmunt Bauman (2006) would say, to exploit and explain the anxieties of globalization phenomena as it is built by networking societies. "Fake news" deals with revelations (-gates) and secrets (-leaks) that lead to suspicion about colluding institutions, medias and the initiated few. The "Donald Cook" story is a case in point:

Box 1: The "Donald Cook" Story

> In 2014, a Russian satiric website posted a "fake news" story about the American warship, USS Donald Cook, having trouble with a Russian fighter aircraft in the Baltic Sea. The story read like a letter from an American sailor, Johnny, to his girlfriend Mary. It mentioned the Khibiny, a Russian electronic weapon. This "fake news" was relayed by Facebook in Russian and in English, in 2014, and re-emerged on the 15th of April 2017, on the news programme "Vesti", broadcast on the public channel Rossiya 1, quoting Facebook as the source and ignoring the denials of both Russian and American authorities about the involvement of the Khibiny in the incident. On April 19th, two British tabloids (The Sun and The Daily Star) picked up the story. Then, the American TV channel Fox News used the story to denounce Russian propaganda but without mentioning that it was a fake. The New York Times eventually exposed the fake story, which led to its withdrawal from the site Foxnews.com without comment or excuse.

Source: Les Echos 18/06/2017

Elements of parody, plot theory and post-cold war imaginary feed on each other. The social media and the mass media relay each other, both in propagation and correction, without necessarily reaching the audience they target. The actors implicated are fake-mongers motivated by fun (a satiric site whose prank is relayed at face value), by ideology (an official Russian television station tries to embarrass the American camp) and by money (to benefit from advertising and increased traffic). All this takes place within a "transverse effect": trans-media (from website to television to printed press), trans-lingual (from Russian to English), trans-time (from 2014 to 2017) and trans-border context (from Russia to the UK to the U.S.). This transverse effect, empowered by the Internet global infrastructure and the digital layer of protocols, applications and social networks, disrupts borders and national sovereignty.

In all cases, three major distortions of competition characterize such "fake news": the tension between mass media and social media, the tension between democratic states and authoritarian regimes, the tension between the culture of control of "the elite" (the wealthy decision-makers, including mass and social managers) and the culture of control of "the base" (the basic members of civil society, including many social media users). It is little wonder that the name of Donald Trump should be associated with "fake news", as he plays upon each of these tensions in turn, ideology trumping policy with authority, fun trumping truth while crowd-pleasing, and money trumping mass media but cajoling social media.

2. SPECIFICITY OF DIGITAL DISINFORMATION AND INTRINSIC MECHANISMS

KEY FINDINGS

- Three major distortions of competition characterize "fake news": the tension between mass media and social media, the tension between democratic States and authoritarian regimes, the tension between the culture of control of "the elite" (mass and social media owners) and the culture of control of "the base" (users generating contents).
- "Fake news" is characterized by three major mechanisms: advertising, virality and automation. Combined, these create a threshold effect, from a quantitative augmentation (more advertising, more traffic) to a qualitative phenomenon (unprecedented virality and automation).
- The combination of advertising, virality and automation calls for a new definition to fit this paradigm change, "malinformation" that recombines human *mal*evolence, marketing *mal*practice and engineered *mal*ware.
- The mechanisms of malinformation can be manipulated by all sorts of actors, with various motivations, in all fields of the economy and society. The actors implicated are fake-mongers motivated by fun (to increase their reach and to monetize their influence), by ideology (to embarrass one camp and promote contrary views) or by money (to benefit from advertising).

2.1 The disinformation trap-situation

In this disruptive context, "fake news" is a new phenomenon, not akin to traditional rumour nor is it a new form of propaganda (Badouard, 2017). It can disturb the established mass media "information contract" according to which news is built on facts reported objectively and associated with reliable indicators (sources, quotes, experts). Such an information contract was not free from manipulation, influence and the propaganda purposes of governments or other powerful bodies, but it was framed by professional principles and codes of ethics. "Fake news" can form a new information contract, with social media trumping mass media with its own indicators of authenticity (likes, ranks, friends), augmented by user-generated content and comments. This new information contract supersedes the first and shifts the users' attention from information value to sharing value. It builds on the functions of social media platforms (YouTube, Facebook, Twitter, Instagram…) and their companion search engines (Google, Yahoo!, Bing...) to make content findable and contact feasible. The professionals of "the elite" are now side by side with new entrants, the online communities of users from "the base", some of whom act as amateur or semi-professional contributors or curators of content. As a result, verified facts stand on the same footing as authenticated fakes that mix true and untrue content that has been viralized by robots as well as humans.

The "fake news" phenomenon was subterranean for a while, since the "social and smart turn" of Web 2.0 (2004-2006) that signals the birth of Facebook, YouTube and Twitter. These new entrants were pure players, born from the digital and not from the pre-digital media of the Web 1.0 era (print and audio-visual media having migrated online). They first focused on information as documents (files, attachments, hyperlinks…) and information as data (codes, algorithms…). It took them some time to realize that information such as news was a valuable product to attract and retain audiences. Bringing together these three cultures of information (news, documents and data) is what enabled the formidable business model these digital giants have generated, and monetized with advertising.

This has been apparent since the take-off of their revenues and market value from the moment major social media (Facebook and Twitter) entered the Wall Street market (2012-2013) (see table 1). Without necessary causality, this is also the moment when "fake news" started emerging in the public discourse, around the U.S. elections of 2012 when President Obama achieved a second mandate. Such discourse peaked during and after the 2016 U.S. elections (Allcott and Gentzkow, 2017).

Table 1: Social Media Revenue

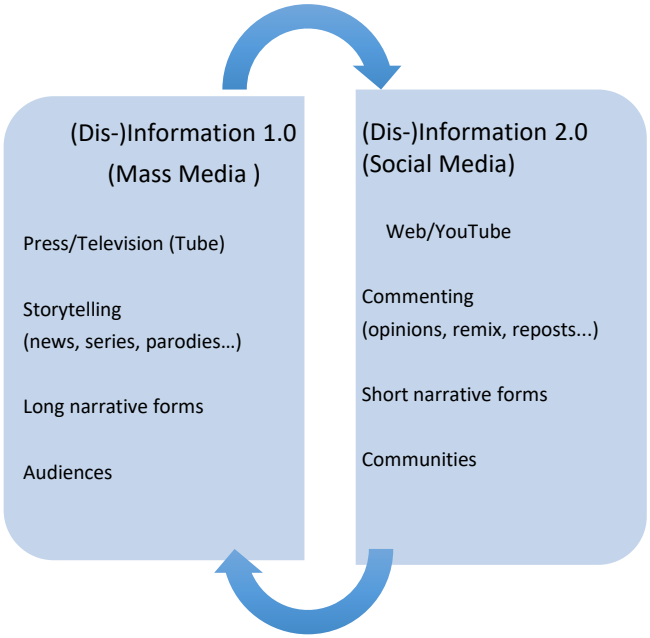
Social Medias	2011	2012	2013	2017
Google	27.72	32.73	38.13	95.00
Facebook	3.15	4.28	6.74	33.70
Twitter	0.14	0.29	0.59	3.26

Source : Compilation from e-Marketer.com via ZDNet.fr

This business model of the digital giants pushes advertising revenue to heights and concentrations that pre-digital media never achieved. It combines the way media generate data and documents with the way data and documents influence media. This interaction is "social" in that it makes monetization possible: the visitors on news websites can be traced and profiled and the profiles bought and sold after a certain number of clicks-per-view, which allows social media to sell adverts and reward users of "the base" for generating content. This interaction is "smart" in that it can make use of automation: algorithm-powered robots can mine the data and detect trends. Then social media can personalize content and target specific users with specific goods and services using cognitive biases and sentiment analysis.

As information 1.0 moved mass media online, information 2.0 created social media augmented with mass media (which still retain the biggest presence for online news). They feedback on each other: Media 1.0 still provide strong long narrative stories and rely on audiences; Media 2.0 provide rapid short narrative comments and rely on communities. "Fake news" creates a trap-situation because though it relies on the same tools and conduits used by verified news to propagate disinformation, as exemplified by the "Donald Cook" story (see figure 1), it benefits from additional advantages.

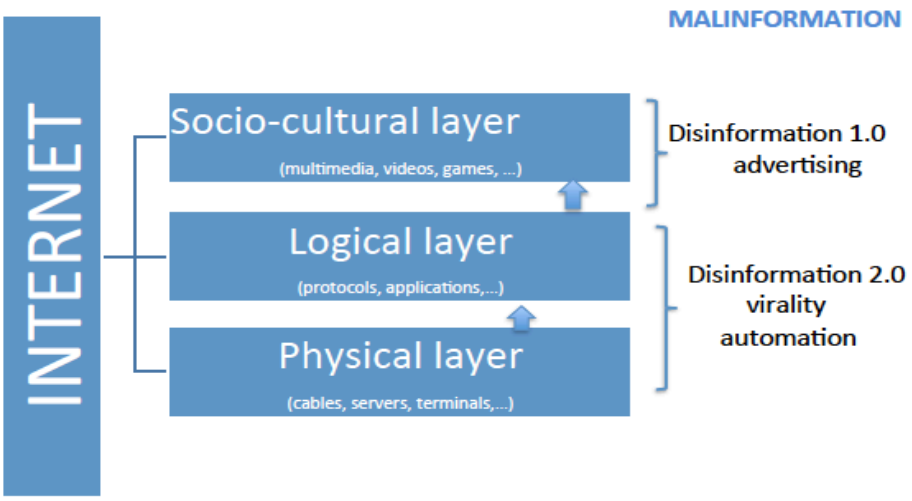
Figure 1: The (dis-)information ecosystem



Source: Frau-Meigs

As disinformation 1.0 travels via the content layer of the Internet, disinformation 2.0 travels via its physical and logical layers (see figure 2). Where traditional propaganda or rumours were relatively contained by scarcity of media, national borders and linguistic barriers, digital disinformation has been unfettered by social media abundance across borders. This enables more automated data mining and more options for viral amplification of content.

Figure 2: The three layers of malinformation



Source: Frau-Meigs

2.2. From "fake news" to malinformation

In combination, advertising, virality and automation generate a threshold effect: what, on the surface, looks like a arithmetical augmentation (more advertising, more traffic) becomes, below the surface, an exponential augmentation and a qualitative change in the nature of news and its effects (unheard of virality and automaticity). This changed relationship between news and public calls for a new definition to fit what is a new paradigm: "malinformation" (Wardle and Derakhshan, 2017; Malwick and Lewis, 2017).

Malinformation is not disinformation as usual. It recombines, in a systemic, smart and social manner, disinformation 1.0 by mass media with human agents, and disinformation 2.0 by social media with human and non-human agents. It carries with it several notions of "bad", "ill" and "wrongful" implicit in the prefix "mal-": human *mal*evolence, marketing *mal*practice and engineered *mal*ware (Frau-Meigs, 2019).

The on-going definitions, most of them proposed by journalists or fact-checkers, all suggest replacing "fake news" with the term disinformation, as in the case of European Journalism Network. Some working definitions rely on typologies, like the Decodex (created by Le Monde): clickbait, satire or parody (referring to material reposted at face value), erroneous information, rumours (unfounded information), plot theories, manipulated information and data (taken out of original context). [1] Useful as they are, these terms and expressions all underestimate the role of "the base", including fake bases powered by artificial intelligence. They tend to look at the production part of malinformation, not at the reception and propagation part. The EU definition provided by the High level group on online disinformation comes closer to a multi-layered and comprehensive definition, as it recognizes a "risk of harm" and the role of "different audiences and communities": "The risk of harm includes threats to democratic political processes and values, which can specifically target a variety of sectors, such as health, science, education, finance, consumer choice and more. It is driven by the production and promotion of disinformation for economic gains or for political or ideological goals, but can be exacerbated by how different audiences and communities receive, engage, and amplify disinformation." (2018, p. 4).

Capturing the complexity of malinformation implies considering facts that are blended with fabricated forgeries, as well as user practices fused with non-human automated fake accounts and comments, fabricated crowds and zombie networks. The digital behaviour of "the base", any base, be it a political group or enthusiasts of a game or product or activity, with its possibilities for participation-contribution-republication, creates unprecedented amplification, supported by the three layers of the Internet, including the infrastructure of the physical layer. Malinformation goes beyond the surface understanding of the phenomenon as partisan political debate or poor journalism (Fletcher and Nielsen, 2017). It goes deep into more pernicious and toxic uses of artificial intelligence protocols, platforms and ontologies to undermine State sovereignty and information integrity. It builds not just on one dimension of information culture, news, but also on documents (that can be morphed, mixed, memed…) and data (that can be aggregated, altered, adulterated…).

It points to a mixture of low-cost news production and high precision data mining and laundering.

[1] https://www.lemonde.fr/verification/

2.3. Major mechanisms

The traditionally disjointed spaces and practices of parody and propaganda can be disruptively re-joined via social media and big data. Independently of their content, the modes of production and of circulation of information have modified the hierarchy of information setting the priorities and the agenda of news. They have altered its mechanisms of added value replacing verifiable truth with the value of sharing. Three major mechanisms are at work, advertising, virality and automation, that all exploit human motivations as re-engineered by neurosciences and artificial intelligence. These smart mechanisms range from influence by distortion and counterfeit with relatively benign consequences to hacks and leaks designed to interfere with countries' sovereignty. While passing themselves off as neutral or benevolent sources they yet have damaging consequences.

Advertising can be directly connected to malinformation disseminated via social media because of a business model that relies on algorithms that do not take into account accuracy and objectivity of information sites but rather focuses on engagement with stories (verified or fake) and the ranking those stories automatically generate. Such engagement is the gauge of attention and becomes monetized, either by selling engaged users to private sector sponsors or selling those same engaged users to content providers of the base increasing their huge following. The e-repertoire of strategies that magnifies this process relies not only on ranking algorithms but also on influencers, clickbait and psychometric targeting. Fake influencers can produce real recommendations for products from fake accounts, as in the case of Louise Delage on Instagram in 2016.[2] Clickbait sites such as OhMyBuzz or Woocom are numerous and manage pools of "fans" for sale. Click detection and sentiment analysis can also target micro-sections of the population with specific messages (Kroll et al., 2016).

Virality also powers malinformation with its own e-repertoire of strategies, such as echo chambers, filter bubbles, memes and trolls that can lead to astroturfing, with the additional "transverse effect" of trans-border, trans-lingual and trans-media virulence of networks as exemplified by the Donald Cook story. They contribute to the multiplier effect of catchy headlines and cascades of retweets and reposts. Echo chambers are affinity spaces where rumours can be amplified while filter bubbles are personal information ecosystems that automatically push certain types of information to users and isolate them from pluralistic contents (Pariser, 2011; Borgesius et al., 2016). Memes, basic units of information such as jingles or logos, can be viralized via hashtags and depositories or data banks to use freely and indiscriminately, as with the case of Pepe the frog, that became an alter ego for Trump during the 2016 elections (before it moved to France!).[3] Trolls are humans that are organized in "farms" to push partisan or catchy content that can polarize communities, as in the case of the whole village of Veles in Macedonia where young people made as much as 5 000 dollars a month clicking content favourable to Donald Trump.[4] They can be part of astroturfing that simulate a whole online militancy and spread the fake feeling that many people support a minority idea or sentiment (Kalogeropoulos et al, 2017).

[2] https://www.theguardian.com/technology/2016/oct/06/shell-drink-to-that-fake-instagram-louise-delage-profile-highlights-alcoholism

[3] https://www.lesinrocks.com/2017/01/21/actualite/de-pepe-the-frog-a-pepe-pen-grenouille-devenue-embleme-de-fachosphere-11903601/

[4] https://www.buzzfeed.com/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo?utm_term=.xdqE1v4MKP#.wh4kjQVr8X

Automation also buttresses malinformation with crowdturfing, social bots, zombie networks and all sorts of artificial intelligence-driven malware (viruses, worms…) and adware (spam). Crowdturfing is like astroturfing, except that its fake militancy is machine generated to attack e-commerce platforms that function on crowdsourcing like Amazon and TripAdvisor. Social bots are opinion-faking programmes that can repost "fake news" at determined intervals of time and have the capacity to link onto feedback loops that augment polarization. Botnets or zombie networks can be computers commandeered to execute malevolent cyber-attacks at a distance, either by altering the usual function of targets or by stealing targets' data or funds. Viruses have the capacity to infiltrate systems by a backdoor or take advantage of security flaws whereas aggressive spams can be implanted in navigators and redirect advertising towards sites such as DomaIQ/SoftPulse or Amonetize.

The combination of these three mechanisms leads to the repertoires of e-strategies specific to malinformation. They all borrow the same circuits as information. They make it difficult to establish malinformation as illegal or illegitimate. Finally they put the burden of proof on actors that are not necessarily trained to detect it (Vargo et al., 2017).

3. SOCIETAL COSTS

KEY FINDINGS

- The convergence of mass media and social media creates a combination of interests that can create harm to the integrity of information. The cost weighs on the profession of journalism, especially as it stimulates the monetization of "fake news" for traffic and concentrates advertising revenues in the hands of a de facto U.S. duopoly, Google and Facebook, with far-reaching consequences for the EU Digital Single Market.
- Risk of harm to the integrity of elections can be due to hybrid threats. Such hybrid threats are part of "information cyberwarfare" campaigns from foreign actors, supported or not by domestic actors, which can lead to the polarization of some segments of the population of one given country as a result of phased, drip strategies that may either polarize voters or disenfranchise them.
- Co-regulation and re-intermediation initiatives pave the way for new ways of dealing with information disorders, such as public interest alliances, or initiatives that are inclusive of "the base" so that it does not feel separated from society and sees its online practices included as part of the solution, not the problem.

The mechanisms of malinformation can be used and manipulated by all sorts of actors, with various motivations, in all fields of the economy and society. Some stand to benefit, others to lose (Wardle and Derakhshan, 2017). In any case, whatever the mechanisms involved, human responsibility cannot be delegated to data and algorithms as these are elaborated by humans and embedded with slants, opinions and prejudices that should not be used as arms of social disruption (O'Neil, 2016).

3.1. Fakedom in all sectors

Although there is no full inventory of the threat posed by malinformation, some indications point to the fact that in some countries it has become a business, tantamount to "fake news laundering". In the United States, Jestin Coler is known as the "Fake news King" with his Disinfomedia company that operates sites similar to mass media organisations, like USAToday.com.co, WashingtonPost.com.co or DenverGuardian.com. Another entrepreneur in fake news mongering is Paul Horner who operated sites like News Examiner, CNN.com.de, CBSnews.com.co and ABCnews.com.co. Coler and Horner claimed advertising revenues around 10 000 dollars to 30 000 dollars per month.

The European market is less susceptible to such entrepreneurship for the moment as advertising practices are more under control, but nonetheless some research indicates the extent to which weaponizing "fake news" can damage trust. The world of finance can be hit, as in the case of hoaxcrashes that can hurt companies by devaluing them. Thierry Berthier looked at six major hoaxcrashes in France and determined several motivations for the attackers: political activism (protest groups in favour of the environmental protection or over the Syrian conflict), and economic profit (speculative investments connected to fake puff news, or short selling connected to brand smearing). He insists on the rapidity of the attack as well as the volatility before detection.

The example of Vinci shows that after publication by Bloomberg of a "fake news" (for short selling) about an accounting mistake in November 2016, the share dropped by 18,28% in a few minutes. The group lost (momentarily) 7 billion euros (Berthier, 2017).

The market at large can also be hit by fake name brand sites and fake name brand suppliers, a practice that existed before social media but has been hugely facilitated by its connection with e-commerce and the transverse effect of malinformation.

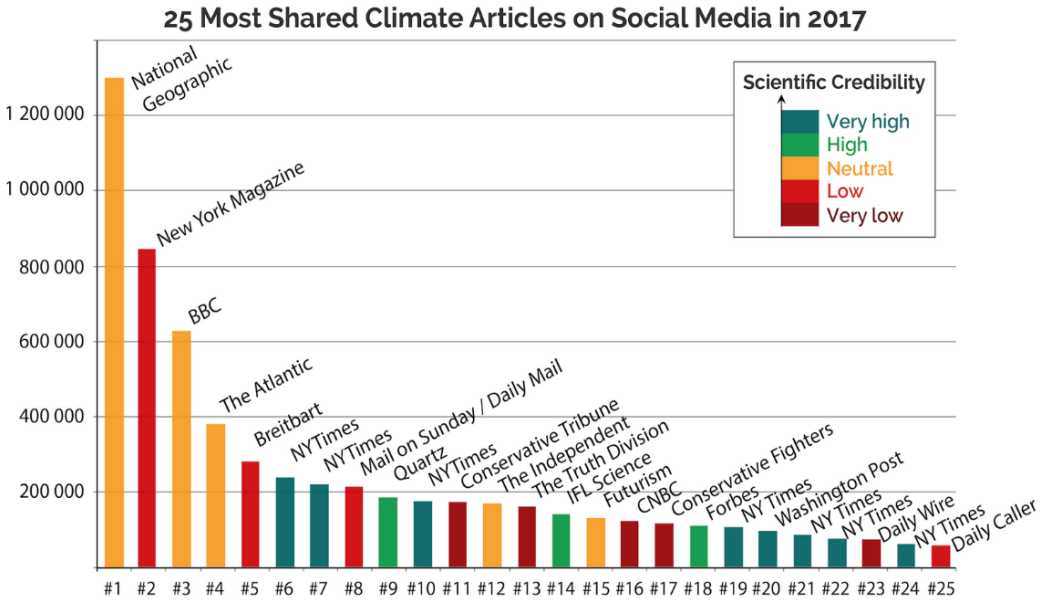
In China, for instance, Alibaba has advertised a showroom where people choose from fake brands (alibaba.com/showroom/fake-name-brands.html). An article by Alexis Madrigal in The Atlantic (January 10th 2018) traces how this practice can be amplified by influencers in Ireland using sites like Shopify and applications like Oberlo to mount a fake company without any overhead or inventory but able to deliver fake goods via AliExpress, a subsidiary of Alibaba, a three-pronged practice known as "dropshipping".[5]

The health sector is among the most targeted by "fake news". Fake-vax can have real life consequences, as in the case in Italy where a nurse was found to have faked the vaccination over 500 children most likely influenced by a campaign of malinformation that associated vaccination with a rise in autism. The "fake news" in that case emphasized the supposed collusion between the pharmaceutical industry, public health policy-makers and the specialized health magazines, all pushing for unnecessary measures. The non vaccination of those children has been linked to an outbreak of measles with more than a 1 600 cases in the first four months of 2018. Other countries were affected, such as France and Romania. This kind of malinformation is also connected with an augmentation of "science denial", that is most obvious in the plot theories and urban legends associated to climate change, and that can lead to fake data. Climate feedback, a scientific initiative that aims at "sorting fact from fiction in climate change media coverage", published a survey showing that half the articles published in 2017 contain elements of fakeness and were more likely to be shared than the others (see Table 2).[6]

[5] https://www.theatlantic.com/technology/archive/2018/01/the-strange-brands-in-your-instagram-feed/550136/

[6] https://climatefeedback.org/most-popular-climate-change-stories-2017-reviewed-scientists/

Table 2: Most shared climate articles on social media



Source : Climatefeedback, 2018

The vast majority of fake websites and fake accounts tend to rely on votes, comments or bots to generate traffic, which in turn generates revenue. Some traffic can be redirected and rewarded by commissions.

Real and fake news are mixed so as to make it harder to tell one from the other, which is tantamount to news-laundering. As regards democracy, two indispensable social functions are particularly at risk, journalism and elections.

3.2. Cost to journalism and risk of harm to integrity of information

Journalism as a profession and a sector cannot be equated to media, as media ownership, in its trend towards concentration in Europe and worldwide, has become less and less controlled by media professionals and more and more by entrepreneurs from other sectors (with a few exceptions such as Bertelsmann and Axel Springer), like construction (Bouygues), aviation (Lagardère) or finance (Finninvest) (Aalberg et al., 2012).

3.2.1. Concentration of ownership and advertising

Transnational media ownership creates a lot of dependency on US-based social media platforms (see table 3). Since 2016 in the USA, the six companies with highest media revenue are Alphabet/Google, Disney, Comcast, Fox, Facebook, Viacom, CBS and Newscorp; Vivendi and Lagardère in France follow far behind in revenue.

Table 3: The biggest media corporations in the world (according to their annual revenue 2016)

Source: Institut for Medien und KommunikationsPolitik, Statista 2017[7]

Among those corporations, two, Google and Facebook, control most of the advertising manna in a transnational manner, interlocked with mass and social media to the point that in 2017 they controlled 80% of the market in the U.S. and in Europe. By contrast, the advertising revenue of American newspapers had gone down to 16 billion in 2017 (compared to 50 billion in 2006). In the digital world, including mobile phones, Adsense (Google) and Audience Network (Facebook) have invented the online business model based on liberal market competition that does not condone state aid or other types of support. This business model, partly based on the monetization of user-generated content, favours "fake news" as engagement and sentiment driven traffic and ad revenue.

It pushes the sharing value of news rather than the information value of news and has demonetized news while making it one of the driving forces of the social media feeds and lists.

As a matter of fact, there are few incentives online to look for verified facts and even for fact-checking, which makes the production of information very costly for mass media. On the contrary, low cost news, including "fake news", is rewarded for being shared. An example of how this works is offered by YouTube, working with Adsense (same parent company, Alphabet/Google) which paid about 0.8 dollars on average for CPM (1000 views) in 2016. But they do not start paying until users have generated up to 100 dollars (or 70 euros) in revenue. As for YouTubers who are influencers, their revenue can be higher if they are sponsored, in which case they go through MultiChannel Networks (MCN) that can take up from 10% to 40% of the ad revenues they bring in, according to Olivier Duffez. Such procedures may explain why Google and Facebook are often mentioned in the malinformation ecosystem. American legislation provides further advantages to these companies by allowing data aggregation and sales of client lists to third parties, which enables psychometric profiling. Finally they are protected by the Telecommunications law of 1996 that lifted anti-trust laws to facilitate the expansion of the soft power of the digital pure players at home and abroad.

[7] https://de.statista.com/infografik/2342/die-top-10-medienkonzerne-weltweit/

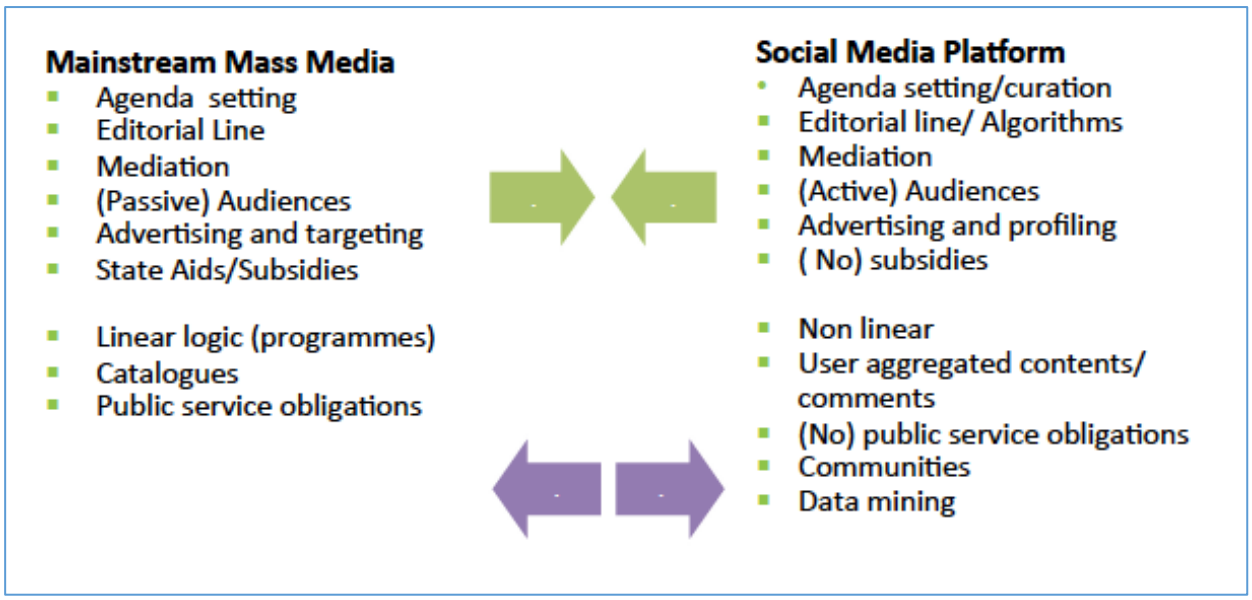
As a result, Google AdWords (on sites owned by Google) represented 71,5% of the total revenue of Google in 2017, while Google AdSense (for partners, including mass media publishers) represents 16,5%. Google gets only 12% of its revenue from sectors outside advertising, according to WebRankInfo.[8] Facebook, via Audience Network, shows a similar behaviour, most of its revenue being generated by advertising, with a clear increase after 2012 and the introduction of advertisement (called 'featured posts') in its "News Feed".

Additionally, the digital giants have become infomediaries: Facebook Instant Articles and Google AMP are portals allowing direct access to mass media, which enables them to extract extra revenue. Google display can boast more than two million websites reaching 90% of the world users. These infomediaries have become necessary news brokers between advertisers and editors and publishers, augmenting the dependency of mass Media 1.0 on Media 2.0. This dependency siphons billions away from the potential added value for creating verified quality information. The share of revenues AdSsense gives back to its affiliates has shrunk over time and represents 71,5% in 2016, according to Parse.ly.[9] Both Google and Facebook lack transparency in that domain.

3.2.2. Convergence between mass media and social media: friends and foes

In their recent evolution, besides capturing advertising, social media platforms have appropriated most of the characteristics of mass media, becoming editors, curators and agenda-setters (see figure 3). Their status as "web hosts" is still protected by article 230 of the Decency Act of 1996 but in terms of perception by users, they function as media.

Figure 3: Two-Stepped convergence of mainstream mass media and social media platforms



Source: Frau-Meigs

[8] https://www.webrankinfo.com/dossiers/google/resultats-financiers

[9] https://blog.parse.ly/post/5194/referral-traffic/

However, in the process mass media and social media act as frenemies, both rivals and friends, co-depending on each other due to the inter-meshed interests they share. Many asymmetries exist not only between mass media and social media but also between social media in the U.S. and in Europe. The European Digital Single Market has not been able to create search engines or social networks of a size large enough to compete with the American ones. The American players are in a controlling market position and do not allow many new competitors to emerge. They are not subjected to the same tax obligations as European media. They are not obligated to accept the same civic responsibilities. They are not obliged to provide transparent information about their sponsors and clients, which may impact referencing and ranking. They are not constrained by the same advertising rules as mass media (during election times for instance, or related to product placement…). They do not have public service obligations like promoting diversity, pluralism, and protection of minors. They are under no compulsory requirement to apply codes of good conduct. Rather, their own "terms of service" are supposedly established with their clients, a relationship which reinforces the alliance between the platforms and their user "base". The e-commerce directive even shelters them with "liability exemptions" for the content they distribute, requiring them only to respect "notice and take down" rules (in relation to hate speech and terrorism, considered as crimes in Europe). Therefore, nothing hinders American-based social media from spreading malinformation for traffic and profit. Thus smart action trumps the smug inaction of the European Digital Single Market with its belief that a quality information product will prevail (Media Pluralism Monitor, 2016; Newman et al., 2017).

Quality information risks harm through these asymmetries as the quality product becomes extremely costly to produce. Meanwhile monetization of malinformation attracts new journalist entrants with neither training nor ethics and no risk of professional consequences to the field competing with journalists of more traditional mass media. Journalists in reference media are under pressure to double check their verified news, which increases the costs and adds time delays. And they do so in a context of reduced editorial staff and increased marketing staff of large conglomerates putting constant pressure to perform for their shareholders (Cagé et al., 2017).

3.3. Cost to integrity of elections and hybrid threats

Advertising and elections are closely related, especially in the context of American mass and social media. The relationship between advertising and elections has different implications for other countries. In the United States, advertising is protected by the first amendment (1976), and this covers political advertising including negative and ad hominem ads. Since 2013, the internet click "Like" is protected as well. Furthermore, elections in the U.S. are covered by laws that have been modified (Citizens United, 2010), allowing corporations and unions to make unlimited contributions to candidates or parties in federal elections. Political action committees (PACs) can pool funds from individuals, corporations or unions without specific ceiling in their support of political parties or candidates.

These PACs are section 527 of the U.S. Internal Revenue Code associations that are exempt of taxation, which has opened up the sluices for political advertising. With the transverse effect of malinformation, this provides potential for alien interference, even if the law forbids foreign financial funding in federal elections (1971).

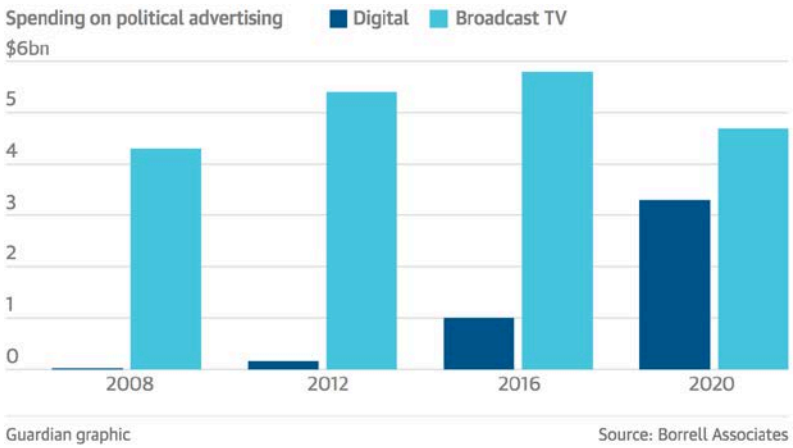
3.3.1. Public relations vs. journalism

Such a situation of "free market of ideas", according to the metaphor first expressed by Justice Judge Oliver Wendell Holmes, Jr. in 1919, as "the competition of the market" and first expressed in a United States Supreme Court decision 1953, by Justice William O. Douglas, favours the transfer of commercial marketing smart methods to political marketing.

The financial manna around the elections business has led to the increase of huge public relations firms and lobbying companies. The United States finds itself in a situation close to that of the 1920s, where the borders where so porous between journalism and public relations that the profession had to re-invent itself around journalism schools that taught professional guidelines such as the norm of objectivity and commercial independence (via advertising and subscription).

PR firms, with political storytelling and spin doctoring, are in direct competition with mass media companies to capture the advertising revenues generated by campaigns. The 2016 elections generated a record 9,8 billion USD, half of which went to television alone. PR firms are gradually realizing the potential and reach of social media and spent 14.4% of their budgets in them in 2016, according to Borrell Associates.[10] Facebook was the prime destination of this money, partly because of its mobile extensions (see Table 4).

Table 4: Spending on political advertising in USA (2008-20)



Source: Borrell Associates

Such a strategy reinforces the interdependence of media 1.0 and 2.0 and fuels malinformation: mass media 1.0 broadcast political storytelling and social media 2.0 relay it, while producing their own that in turn is picked up by mass media, in a "fake news laundering" loop. Firms such as Cambridge Analytica benefited from this mechanism and used information as documents and as data to profile electors for their clients, worldwide.

Donald Trump's chief campaign strategist and business-man Steve Bannon, also creator of Breitbart News with its alt-right vision and a co-founder with banker Robert Mercer of Cambridge Analytica, is said to have used its services for micro-targeting. Such a strategy maximizes transverse effects, and facilitates access of foreign actors into the system permitting the manipulation of online communities to subvert election results (Bartlett, 2017; Huchon, 2017). Facebook told the U.S. Senate in 2018 that disinformation sources related to Russia paid 170 000 USD to Facebook for political ads between June 2015 and May 2017.

[10] https://www.borrellassociates.com/industry-papers/papers/political-advertising-outlook-2014-and-beyond-jul-14-detail

3.3.2. Vulnerability of elections

Elections are particularly vulnerable times in democratic countries because they are priming moments when citizens update their opinions and values. Elections provide a privileged moment for focusing on the future of a country, with a natural polarization due to voting outcomes (yes/no; in/out). Such conditions make electoral campaigns rife with malinformation and its manipulative strategies that go beyond propaganda to create liquid fears to stir divisive cultural memories. Such strategies can be intimidatingly effective in disenfranchising people, making them believe that their votes do not count or that they are not legitimate voters, all of which can modify the outcome of the vote without changing any voters opinion of an issue or a candidate (Huighe, 2016).

Among the strategies for malinformation, voter suppression is common. Wrong or misleading information can be sent to mislead potential voters as to where and when the vote takes place. This happened during the 2016 elections, with fake SMS messages inciting people to vote in advance for Hillary Clinton. Rumours of pirated voting booths or modified voting registrars are also used to demotivate people from voting. Astroturfing and crowdturfing are mechanisms that can be implemented to polarize online communities. A fake militancy can make citizens believe that their minority opinion is shared by many. Hyper-partisan pages, driven by trolls, can make use of echo chambers to propagate fallacious information that has enough of a kernel of truth in it to cast doubts about candidates or to create sensation (revelations, secret affairs and hidden bank accounts…) (Rožukalne, 2015).

More insidious, drip campaigns, that start much earlier than the actual election period, can be orchestrated and machine-driven, with infiltration by foreign or domestic actors wanting to bias the agenda before it is even set. The case of United States of America vs. Internet Research Agency (IRA) in 2017, put forth by the grand jury of the District of Columbia, revealed that thirteen Russians and three Russian entities had used such a strategy to interfere in the U.S. elections. This whole range of malinformation has been set in the category of information warfare and hybrid threats by NATO and the Council of Europe and has led to the creation of a Stratcom Center of excellence that monitors such drip strategies in regions of Europe that are susceptible to it, such as Lithuania and Latvia (Stratcom, 2017) whose geographic and demographic position makes them vulnerable to Russian interference. Each country has called on an "army of elves", to keep the armies of trolls existing elsewhere in check.

Within the European context, the repertoire of drip strategies and their various phases of malinformation, can be detected through several examples that have attracted public attention, such as the Cambridge Analytica case. They recombine different mechanisms of malinformation that lead to hybrid threats:

- drip dissemination of topics to create viral trends around polarizing topics well before the election time;
- hack of strategic data to cause surprise or panic at determined stages of the campaign;
- whitewashing or camouflaging "fake news" via social media to amplify controversy via mass media and conversely ;
- secret collusion or real objective complicity or alignment of interests with extremist or minority groups in a country allowing internal and external actors to buttress each other.

The asymmetries between Media 1.0 and Media 2.0 can cause risks of harm to information about elections.

Asymmetries illustrate the tension and competition of models between democracies that allow free elections and authoritarian regimes that pass themselves off as democratic because they allow elections but control those and promote dictatorial political leadership in exchange for degrees of market freedom.

Malinformation affects not only politics but also diplomacy and defence sectors and puts the culture of peace at risk, because it creates a situation of cyber war. This war results in cyber security issues that can escalate into a cyber arms race and challenges the neutrality and the global infrastructure of the Internet. The risk of splitting internet into various Internets is very real. China has created its own alternative cyber ecosystem and can spread its computational propaganda to neighbouring targets while remaining protected by its digital great firewall (King et al, 2016).

In the end, the issues raised by integrity of elections are like those raised by integrity of information, in the ecosystem of malinformation. Not relying on information is cognitively damaging because it inhibits decision-making, be it political (voting), economic (buying, selling), health related (vaccinating), or some other essential human and social activity. The threats extend not only to news as produced by journalists but to information as provided by researchers and all other knowledge-producers.

The stakes relate to transparency, trust and reliability. The direct impact of malinformation, though not proven yet by the nascent research protocols, may lead to long-term counter veiling deterrence measures which will reduce freedom of speech and freedom to receive and impart information, with creating damages on democratic processes. The democratic cost, hard to prove or quantify because it is dissipated and often indirect, is lack of trust in institutions and media (OSCE, 2017).

4. RESPONSES: FROM SOFT TO HARD TO SMART

KEY FINDINGS

- In terms of self-regulation, fact-checking has emerged as the choice response of the private sector. It can be performed by humans such as journalists, activists, web-moderators and even users and "the base". Machine-learning and automation can fact-check too.
- The risk with such self-regulatory measures and other automated solutions such as deplatforming, shadowbanning and other means of taking down fake sites and accounts and zombie networks is that governments outsource the policing of the Internet to the private sector and to private actors. Censorship is then undertaken by platforms that can also use it to optimize their own activities and increase the range of their "services". The criteria for demonetizing accounts, for preventing access are not clear and transparent and touch the limits of self-regulation where the private sector is both judge seeking to filter and partisan interested in amplifying.
- In terms of regulation, transparency is the main goal. Existing regulation demonstrate this in the realms of advertising, sponsored content, funding of electoral campaigns or in the processes leading to elections and to the actions of political parties and other actors, domestic or foreign.
- The risk of regulatory measures is two-fold: censorship and surveillance, both abundantly evident under non-democratic regimes. Censorship can be operated directly by States, closing down sites or jailing journalists and opponents. Surveillance initiatives can monitor systems and black-list journalists, bloggers or whistle-blowers. The negative and unintended consequences of any regulation need to be weighed carefully in democracies and markets where fundamental freedoms and principles of freedom of expression and privacy support trust.
- In terms of education, media and information literacy (MIL) can help reduce the attraction of the storytelling methods of "fake news" and propose attractive counter-narratives in response, using critical thinking thus reducing the asymmetrical advantages of cognitive biases on which malinformation relies.
- MIL should be part of the core curriculum of young people in schools, as one of the key competences to deal with XXI century information disorders and digital citizenship. Placing it among the requisite competences to be evaluated by PISA would also provide for more visibility, legitimacy and assessment, thus resolving certain chronic vulnerabilities of MIL.

Malinformation can be countered at two stages in the amplification flow: at the source before the event, when "fake news" is produced and later, when it is disseminated for consumption. The types of responses and of actors are not the same. At the source, self-regulation is the preferred action of the private sector of mass media and social media alike while regulation is the temptation of the public sector. At the mouth of the river of information, education is the favourite solution of the civic sector and the users.

4.1 Self-regulation: fact-checking on the rise

Besides the traditional media accountability systems (MAS) that mass media have called upon to address this issue (handbooks, norms, curricular changes, press councils…), a new solution has taken centre stage, fact-checking. The Poynter Institute for Media Studies has even created a norm for excellence in fact-checking for the new international network of fact-checkers (IFCN). Though it started after 9/11 in the USA, fact-checking has developed considerably during and after the various elections of 2016-17, with many national and transnational initiatives put in place following the transverse effect of malinformation.

Box 2: Examples of international initiatives in fact-checking

Citizenevidence.org (UK): Amnesty International with YouTube Data Viewer

Crosscheck (USA): First draft projects for election monitoring

FactCheckEU: European crowdsourced platform for users

Faktabaari (Finland): verification by journalists for elections

Firstdraftnews. com (USA): Harvard University project, to coordinate verification with research

Full Idea Project (Latvia): Baltic Centre for Media Excellence

Reality Check (UK): BBC program

RevEye (USA): Google plug-in to verify images

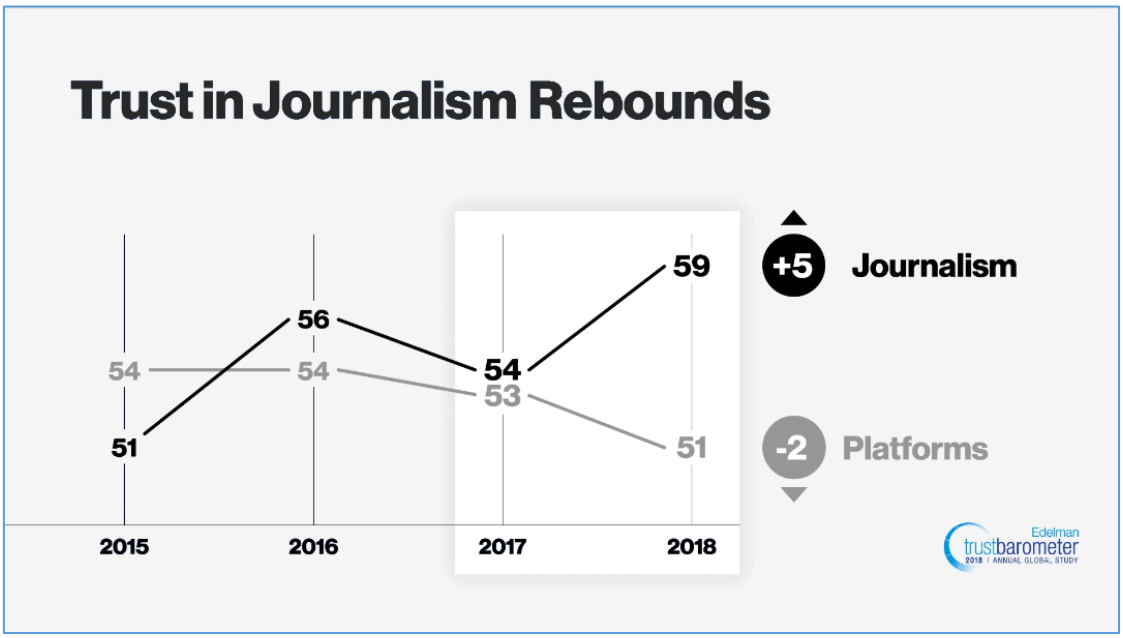
TinEye (Canada): website to check images

TweetCred (USA): plug-in with credibility scores provided by crowdsourcing

Source: Frau-Meigs

Political fact-checking is practiced by journalists and other actors creating potentially positive tension between internal processes of the profession and external public contradictory processes. Fact-checkers can be journalists but also activists, researchers or politicians. There is a wide range of devices and practices around the world. The report by COMPACT, an H2020 research project published in February 2018, points to both advantages and limitations of fact-checking. It stresses the fragility of funding and the lack of resources as well as their difficulty of getting their results to the general public. A lack of methodological transparency has also cast some doubts as to neutrality and has led many conspiracy websites to undermine fact-checking. However, it is to be noted, without presumption of a direct cause-effect relationship that the level of trust in mass media has risen since 2017, while trust in social media has descended (see table 5).

Table 5: Trust in journalism (2018)



Source: Edelman Trustbarometer, 2018[11]

Social media platforms have also been supporting and sponsoring fact-checking. Besides funding projects like FirstDraft or Crosscheck, they have used machine learning to push for automated fact-checking. Some solutions use credibility scores and filters such as TweetCred for Twitter. Others favour data annotation and curation, such as the Jigsaw project or the French start up Storyzy (Helberger et al., 2018).

Social media, however, tend to have their own new MAS, mostly around purges of fake accounts, moderation of content and diminution of monetization for users that manage troll farms. They continue their tacit pact with their users of "the base" to promote fact-checking by crowdsourcing, creating alerts and signalling and improving extensions to be installed on navigators like Google for the Decodex of *Le Monde*. Facebook has launched the "News Integrity Initiative", a consortium with 14 million dollars to help internet users create fact-checking projects.

Social media and mass media have also developed some alliances. Twitter, Facebook, Google and Bing are part of the non-partisan "Trust Project" to create indicators and enhance findability of quality news. Google launched the "Digital News Initiative", with 150 million dollars over 3 years to support European projects for mass media to innovate and increase their online presence, especially on mobile phones and applications.

As for the integrity of elections, social platforms have improved their inner systems to ensure they are less vulnerable to cyber-attacks. Facebook and Google each have developed programs to understand and prevent foreign interference.

[11] https://www.edelman.com/trust-barometer/

They have let it be known that they are collaborating with governments, especially commissions in charge of monitoring elections. Facebook has extended its "ThreatExchange program" as a social platform allowing digital companies to share information about cyber-security. Google set up "Project Shield" to address DDoS (Distributed Denial of Service) types of attack. Twitter is moving to a new functionality, coding private messages, much the way dark social services such as Telegram and WhatsApp do.

The risk with such measures and other automated solutions such as deplatforming, shadowbanning that take down fake sites and accounts and zombie networks is that governments outsource the policing of the Internet to the private sector and to private actors. The censorship is then done by platforms that can also use it to optimize their own activities and increase the range of their "services". The criteria for demonetizing accounts or for preventing access are nor clear neither transparent, and touch the limits of self-regulation where the private sector is both, a judge (filtering) and a party (amplifying).

4.2. Regulation: towards a transparency agenda

The democratic States have taken some steps to nip malinformation in the bud. They address its three mechanisms by asphyxiating advertising, slowing down virality and retro-engineering automaticity. They nonetheless have to consider the negative and unintended consequences of any regulation in a democracy where fundamental freedoms and principles of freedom of expression and privacy prevail and in fact foster trust from "the base".

Some negative practices have been put in place by non-democratic regimes that can be seen as red flags for malinformation. They fall into two categories: censorship and surveillance. Censorship can be operated directly by States, closing down sites or jailing journalists and opponents. The case that has attracted the most attention is the Rappler takedown in the Philippines in 2016. This website denounced more than 300 "fake news" propagating sites and was taken down by President Rodrigo Duterte in an attempt to stifle political opposition during elections. Other authoritarian regimes have put in place monitoring systems that are less obvious, such as black-listing bloggers or whistle-blowers.

In the European Union, some governments have moved on to hard law and regulation. Besides some measures by the UK, Sweden or Italy, Germany and France are the two countries proposing the most comprehensive attempts at legislation. Germany has opted for a law on social media, known as "NetzDG". Implemented in January 2018, it compels the social media platforms to take down illegal content within 24 hours and subjects them to substantial fines if they do not comply (up to 50 million euros). It does not provide for users to appeal decisions and has been criticized as dangerous for political expression. Facebook messages by Beatrix Von Storch, vice-president of The Alternative for Germany party were erroneously taken down in the very first few days of implementation.

In spite of heavy criticism, France has also voted in November 2018 a law "to fight against the manipulation of information". It purports to make the social media platforms remove content within 48 hours and obliges them to comply with rules of transparency of sponsored content. It authorizes the High authority on audiovisual media (CSA) to suspend channels that are suspected of foreign influence. There is also a procedure for rapid suspension of content, especially during elections, requiring judges to stop the spread of "fake news". This measure seems the hardest to implement as it leaves the burden of proof in the hands of the judges, a task they are reluctant to carry out for practical and judicial reasons.

However, these measures all ignore the "transverse effect" of malinformation, besides not taking into account the very protected status of social media under U.S. law. Hence various initiatives have been started at the level of the European Union.

To asphyxiate advertising, the EU can look at the e-commerce directive and try to restore some symmetry to the very unbalanced current situation between mass media and social media. The status of social media needs to be modified or renovated to accommodate the fact that they have become producers and distributors of information and data. Rules fighting monopoly situations should also be called into action. The use of data is also going to benefit from the GDPR which should afford some leverage to aggregation of data and profiling.

To slow down virality, the EU can put in place transmedia and transborder transparency rules for the loyalty of social media (classification, codes of ethics, guidelines…). This kind of measure can be pushed by tax incentives and other positive means. Regulating intellectual property could also slow down users and producers of "fake news" as they could be fined for misuse of content. As for automaticity, more transparency revealing the humans behind such systems can be required. Retro-engineering modification can also help identify the location of fake accounts and troll farms.

As for integrity of information, positive measures can be taken for funding public media services and for supporting the good working conditions of journalists. Media services can also be helped to move into the mobile industry where business models of the future are developing. For integrity of elections, positive measures can be taken to make platforms and political parties alike more responsible and transparent about their activities and their alliances, especially with foreign third parties. Improving the accessibility of correct information during elections should be a top priority.

There is currently some convergence in the different measures that are being considered in the EU within the Digital Single Market framework. They have some pertinence to the realms of soft and hard regulation of malinformation:

- The e-commerce directive requires more due diligence for Google, Facebook and Twitter in terms of notification of fraud and illicit content.
- The directive on audiovisual media services in its revised form includes platforms with audiovisual content and puts them under the rule of media authorities, which may affect YouTube and visual malinformation while promoting a convergent regulation between mass and social media.
- The new directive on copyright, thought controversial, tries to constrain platforms to put in place tools for automatic detection of illegal content.
- The network and information security directive compels operators and platforms to be transparent about cyber threats and to report their security breaches.
- The general data protection regulation affords a number of guaranties and protections to users, in particular their explicit content and their right to portability of data while requiring the platforms to be explicit and exhaustive about the planned use of collected data.

Besides this series of measures that are taken sector by sector, without coordination, the EU has started a series of initiatives directly aimed at "fake news". The EU delivered a communication on April 26th, 2018, following up on an online consultation and a report by the high level expert group on the subject of online disinformation.[12]

[12] https://ec.europa.eu/digital-single-market/en/news/communication-tackling-online-disinformation-european-approach

The main recommendations point to more transparency (in terms of advertising and elections); they support independent fact-checking and media and information literacy; they also set a process for engaging the social media to put in place a code of good practices with several steps to ensure its implementation.

4.3. Media and information literacy: smart sense-making

To address the tension between the culture of control of the elite and the culture of control of the base, media and information literacy (MIL) appears as the solution of choice, from the perspective of all actors. The purpose is to create some resilience to malinformation and to reboot trust and confidence in institutions and in social media. This will be achieved by reducing the attraction of the storytelling engaged in by "fake news" and to propose attractive counter-narratives in response, using critical thinking and reducing the cognitive biases on which malinformation relies. MIL can help detect cognitive biases in oneself and in others and can sensitize to the utility of journalism and the benefits of a healthy democracy with its attendant online freedoms.

MIL has evolved since its pre-digital stages, incorporating the three cultures of information (news, documents and data). It has its own epistemology and its own research questions while benefiting from its own policies in Europe via the Audiovisual Media Services Directive that brought about a series of harmonisation laws in the years 2012-2014. In the EU, it is placed under the auspices of the DG CNECT. However, research shows that MIL suffers from several constraints that hinder its efficacy to counter malinformation:

- the lack of teacher training in initial or continuing stages renders teachers impotent when faced with the (not so) subtle mechanisms of radicalisation and disinformation;
- the lack of visibility of its presence in the curricula and in the sets of competences demanded by ministries of education and/or ministries of culture and communication, even in countries like Finland or Sweden where they are part of the new multi-literacies programmes;
- the absence of collaboration between university sectors that could associate research and practice, civil society activists and academic experts;
- the dysfunctioning governance at the ministerial levels where no co-regulatory mechanisms coordinate the work of operators, educators and researchers;
- the chronic lack of funding and of evaluation of MIL that leads to a persistent weakness in proving its efficacy and transferability beyond the establishment of "good practices" (Frau-Meigs et al, 2017).

Since 2016, with the disruptive rise of "fake news", MIL has benefited from renewed interest. It has benefited from the rise of fact-checking, as journalists and activists have also invested their expertise in schools.

Box 3. Examples of MIL good practices to counter "fake news"

Bad News Game (UK): game to construct "fake news"

What the Fake (Switzeland): quizz to alert about fake contents

Factbar EDU (Finland): journalists in classrooms to teach children how to detect fakes

Info hunter (France): pedagogical tools to decode news

Lie Detectors (Belgium, Germany): journalists in classrooms to teach the detection of fakes

Mind over Media (USA, EU): platform with crowdsourcing strategies to detect propaganda

MOOC DIY MIL (France): massive online open course to build MIL projects with Savoir*Devenir

Source: Frau-Meigs

Social media platforms have also appropriated MIL. Google and Facebook offer MIL projects often confused as digital literacy. Google proposes projects in Europe, via YouTube, like "be internet awesome" (for children 9-11 year old) or "digital citizenship" for cyber safety.[13] However, MIL experts are concerned about the independence of MIL when provided by the private sectors and actors whose mission is not education but branding and marketing.

In the face of such a positive and yet stagnating context, the most viable solution to protect and promote MIL is to make it part of the core curriculum of young people in schools, from the first to the last grades of public education. Making MIL one of the requisite competences to be evaluated by PISA (not to be confused with ICT competences) would provide for more visibility, legitimacy and assessment, thus addressing chronic vulnerabilities of MIL. Schools and universities would then make it part of their teacher training programmes. Research would also benefit from renewed opportunities to test and observe new pedagogies and new responses to malinformation. MIL could thus contribute effectively to the development of digital citizenship education as promoted by the Council of Europe.[14]

[13] www.google.com/safetycenter

[14] https://book.coe.int/eur/en/human-rights-democratic-citizenship-and-interculturalism/7451-digital-citzenship-education-volume-1-overview-and-new-perspectives.html

5. CONCLUSION: TOWARDS A NEW PHASE OF CO-REGULATION AND RE-INTERMEDIATION

The current situation paradoxically reveals the social utility of the "fake news" panic, as it disclosed the magnitude of information disorders, extant and to come, to the general public and decision-makers (Alava et al., 2017). The joint and coordinated activities of various actors point to co-regulation, understood as the convergence of multi-stakeholders around a common issue. They should design solutions that are then carried through by each of them according to their missions and competences. Multi-stakeholderism in democratic countries can foster flexible and agile alliances and coalitions that include transparency and accountability mechanisms, which in turn can generate trust (UN IGF, 2017).

In the case of malinformation, the actors implicated are the mass media, the social media platforms, the advertisers, the professional associations (journalists, consumers, researchers, teachers and educators…), the governments and the regulation authorities. A process of co-regulation is at work with the high-level expert group on "fake news and online disinformation" that will continue meeting and evaluating the process beyond the initial meeting time (January-March 2018), according to an agile method adapted to such a multifaceted phenomenon as information disorders.

Co-regulation initiatives have also engaged the co-responsibility of many actors, such as the Deepnews.ai project developed by French journalist Frédéric Filloux that aims at using artificial intelligence to improve automated search capacity for news. Another initiative by the EU H2020 project IN-VID takes fact-checking to the fields of research and schools. Other initiatives rely on civil society actors and users, such as Data Detox by the Mozilla foundation or Wikitribune to stifle "fake news" with an active community akin to the one powering Wikipedia.

Re-intermediation can bolster co-regulation in Europe and the Digital Single Market. Two recent initiatives show the emergence of alliances. The Open Internet Project has been launched by several European digital corporations (Qwant, Axel Springer, LeGuide.com, Sinhorcat…) to present a common front to the dominant position of the GAFAM in the European market. The Skyline alliance aims at recuperating some of the advertising revenue that exists in Europe. TF1 (France) has joined ProSiebenSat.1 (Germany), Mediaset (Italy and Spain) and Channel 4 (UK) to create the "European Broadcaster Exchange" to compete with Google and Facebook.

In the long run, it is possible to imagine an alternative system to the strictly commercial system that exists now by adding a "public interest" requirement that reflects more specifically the European heritage and values. Already a whole dorsal structure that could support this system exists, from search engines to social media. Companies carrying them, however, are currently too small to compete credibly with the GAFAMs. There are search engines that do not trace (Qwant), social media that do not post advertisements (Mastodon, Vero…). There are even public algorithms in relation to publicly available application programming interface (API) that provide developers with access to proprietary software applications or web services (Twitter and Yahoo! allow it). Non-commercial initiatives akin to Wikimedia also can foster the information commons.[15]

[15] https://netcommons.eu/

In conclusion, "fake news", and the deeper information disorders it points to, are here to stay. Future trends point to artificial intelligence, augmented reality and other deepfake innovations. There is an urgent need to establish an agenda for research, to help policy-makers, professionals and citizens at large. Adequate responses are not going to come from short tweaks or corrections of the Digital Single Market nor from technology "fixes" by the social media platforms. They are going to come from a collective effort to restore trust in institutions and media and from a collaborative effort to reach out to "the base", with all its contradictions and contributions. Much more inter-disciplinary and inter-cultural research is needed to achieve these safeguards to democracy, and to provide new perspectives on the online public sphere, its opportunities and limitations. The agenda for research should consider all the dimensions of malinformation, from its production to its diffusion and amplification, from its agents to its victims, from its human biases to its technical distortions. A shared understanding of all the implications can ensure that no undue or damaging regulations are applied and that positive, far-reaching propositions are adopted by all.

REFERENCES

- Aalberg, T. and J. Curran (ed) (2012), How Media Inform Democracy: A Comparative Approach, Routledge.
- Allcott, H. and M. Gentzkow (2017), Social Media and Fake News in the 2016 Election, National Bureau of Economic Research, http://www.nber.org/papers/w23089
- Alava, S., D. Frau-Meigs and G. Hassan (2017), Social Media and the Radicalization of youth in the digital era, UNESCO.
- Badouard, R. (2017), Le Désenchantement de l'Internet: désinformation, rumeur et propaganda, FYP Editions.
- Bauman, Z. (2006), Liquid Fear, Polity.
- Berthier, T. (2017), Les 3 F du HoaxCrash : Fausses données, Flash Crash et Forts profits, Chaire de Cyberdéfense & Cybersécurité Saint-Cyr - - IV-Mesure de la Cybermenace - Article IV - 10.
- Borgesius, F. J. Zuiderveen, D.Trilling, J. Möller, B. Bodó, C. H. de Vreese and N. Helberger (2016), Should We Worry about Filter Bubbles? Internet Policy Review, March. https://policyreview.info/articles/analysis/should-we-worry-about-filter-bubbles.
- Cagé, J., N. Hervé and M-L. Viaud (2017), L'information à tout prix, Institut National de l'Audiovisuel.
- Duffez, O. (2018), Comment se faire payer par Youtube? Le guide complet 2018 !, WebRankInfo, 17/01/2018.
- Fletcher, R. and R. K. Nielsen (2017), Are People Incidentally Exposed to News on Social Media? A Comparative Analysis. New Media & Society, August, 1461444817724170. https://doi.org/10.1177/1461444817724170.
- Frau-Meigs, D. (2019), Faut-il avoir peur des "Fake news"?, Documentation Française
- Frau-Meigs, D., I. Velez and J. Flores Michel (ed) (2017), European Public Policies on Media and Information Literacies in Comparative Perspective, Routledge.
- Helberger, N., J. Pierson and T. Poell (2018), Governing Online Platforms: From Contested to Cooperative Responsibility, The Information Society 34 (1):1–14. https://doi.org/10.1080/01972243.2017.1391913.
- Huighe, F-B. (2016), Désinformation : les armes du faux, Armand Colin.
- Kalogeropoulos, A., S. Negredo, I. Picone and R. K. Nielsen (2017), Who Shares and Comments on News?: A Cross-National Comparative Analysis of Online and Social Media Participation. Social Media + Society, October. https://doi.org/10.1177/2056305117735754.
- King, G., J. Pan and M. Roberts (2016), How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, not Engaged Argument, Harvard University, http://gking.harvard.edu/files/gking/files/50c.pdf?m=1463587807
- Kroll, J.A., J. Huey, S. Barocas, E. W. Felten, J. R. Reidenberg, D. G. Robinson and H. Yu (2016), Accountable Algorithms, Univ. of Penn, L. Rev 165 (2016): 1-66.
- 0,Neil, C. (2016), Weapons of Math Destruction: How Gig Data Increases Inequality and Threatens Democracy, Crown Publishers.
- OSCE (2017), Joint Declaration on Freedom of Expression, "Fake News", disinformation and propaganda, March 3rd, https://www.osce.org/fom/302796
- Pariser, E (2011), Filter Bubbles: What the Internet is Hiding from You, Penguin.
- Rožukalne, A. (2015), Internet News about Ukraine and the "Audience Agenda": Topics, Sources and the Audience Aggressiveness, Journalism Research 8 (1):17-37.
- Stratcom (2017), Internet Trolling as a Tool of Hybrid Warfare: The case of Latvia, https://www.stratcomcoe.org/download/file/fid/3353.
- UN IGF Dynamic Coalition on Platform Responsibility (2017), Platform regulations: how platforms are regulated and how they regulate us, http://bibliotecadigital.fgv.br/dspace/handle/10438/19402.

- Vargo, C. J., L. Guo and M. A. Amazeen (2017), The agenda-setting power of fake news: A big data analysis of the online media landscape from 2014 to 2016, New media & society 00 (0):1-22. https://doi.org/10.1177/1461444817712086.

REPORTS

- Marwick, A. and R. Lewis (2017), Media manipulation and disinformation online, Data&Society, https://datasociety.net/output/media-manipulation-and-disinfo-online/
- Independent High Level Expert Group on Fakenews and Online Disinformation (2018), A multi-dimensional approach to disinformation (report), European Commission. https://ec.europa.eu/digital-single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation
- Institut de Recherche Stratégique de l'Ecole militaire (2018), Les manipulations de l'information. Un défi pour nos démocraties https://www.diplomatie.gouv.fr/IMG/pdf/les_manipulations_de_l_information_2__cle04b2b6.pdf
- Media Pluralism Monitor (2016), http://cmpf.eui.eu/media-pluralism-monitor/mpm-2016-results/
- Newman, N., R. Fletcher, A. Kalogeropoulos, D. A. L Levy and R. K. Nielsen (2017), Reuters Institute Digital News Report 2017, Reuters Institute for the Study of Journalism. http://www.digitalnewsreport.org/.
- Pavleska, T., A. Školkay, B. Zankova, N. Ribeiro, A. Bechmann (2018), Performance analysis of fact-checking organizations and initiatives in Europe: a critical overview of online platforms fighting fake news, COMPACT http://compact-media.eu/wp-content/uploads/2018/04/Performance-assessment-of-fact-checking-organizations_A-critical-overiview-Full-Research-1-1.pdf
- Wardle, C. and H. Derakhshan (2017), Information disorder: Toward an interdisciplinary framework for research and policy making (report), Council of Europe Publishing. https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c

DOCUMENTARIES

- Bartlett, J. (2017), The digital guru who helped Donald Trump to the presidency, BBC News, 13 August,

  http://www.bbc.com/news/av/magazine-40852227/the-digital-guru-who-helped-donald-trump-to-the-presidency

- Huchon, T. (2017), Unfair game : comment Trump a manipulé l'Amérique, Spicee, https://www.programme-tv.net/programme/culture-infos/11096267-unfair-game-comment-trump-a-manipule-l-amerique/

This in-depth analysis explores the mechanisms of "fakenews" (redefined as "malinformation") and their societal costs in the digital single market. It describes the risks to integrity of information and to integrity of elections. It highlights the roles of the various actors involved in the production and amplification of such information disorders. Finally, it outlines responses that are being tested in different parts of Europe to deal with the issue while remaining within the human rights framework. The document was provided by Policy Department A at the request of the European Parliament Committee on the Internal Market and Consumer Protection.

PE 626.087
IP/A/IMCO/2018-06

Print ISBN 978-92-846-4124-6 | doi: 10.2861/273718 | QA-05-18-026-EN-C
PDF ISBN 978-92-846-4125-3 | doi: 10.2861/082956 | QA-05-18-026-EN-N