

An inventory of measures to counter “fake news”

25 April 2018



Institute for Information Law (IViR)

Faculty of Law

University of Amsterdam

www.ivir.nl

Authors:

Dr. Tarlach McGonagle

with

Eugénie Coche, LL.B. and Cees Plaizier, LL.M.

and

Melanie Klus, LL.M. (additional contribution)

Contents

Executive summary	3
1. Introduction	5
2. Definitions	8
3. European (legal and policy) framework	15
4. Methods/measures against “fake news”	25
4.1 Preventive or pre-emptive measures	26
4.1.1 Media, information and news literacy	26
4.1.2 Trust-enhancing practices	28
4.1.3 Technological initiatives	29
4.2 Identification and monitoring measures	30
4.2.1 Fact-checking	30
4.2.2 Flagging / labelling / blacklisting	34
4.3 Containing or corrective measures	35
4.3.1 Contextualization of (fake) news	35
4.3.2 Using ‘nudges’ to slow sharing	37
4.3.3 Closing automated accounts	37
4.4 Regulatory and non-regulatory counter measures	38
4.4.1 Regulatory measures	38
4.4.2 Non-regulatory counter measures	44
5. Effectiveness of methods/measures against “fake news”	48
6. Applicability in the Netherlands	59
7. Conclusion	66
8. Bibliography	67
Academic works	67
Regulatory texts	71
European Court of Human Rights: case law	72
Policy documents / miscellaneous	73
News articles, reports and analyses	76
Websites cited	81

Executive summary

“Fake news” has been the focus of a lot of attention in the media and in public debate recently. Against this background and in light of the potential threats of “fake news” for Dutch society, this study has been commissioned by the Ministry for Education, Culture and Science. The aim of this study is to provide an update on the state of the art regarding “methods to counter the spread of fake news”. It answers the following questions: Which methods are internationally known for combating fake news? How effective are those methods? What qualifications can be made about the applicability and relevance of these research results to the Dutch context?

Before answering these questions, the term “fake news” is defined (Chapter 2) and the applicable European (and international) legal and policy framework is briefly sketched (Chapter 3). Given the confusion surrounding the term “fake news”, this study defines it as “information that has been deliberately fabricated and disseminated with the intention to deceive and mislead others into believing falsehoods or doubting verifiable facts; it is disinformation that is presented as, or is likely to be perceived as, news”. The legal instruments playing a core function, which shall be taken into account when developing regulatory measures, are, at European level, *inter alia* Article 10 of the European Convention on Human Rights, Article 11 of the Charter of Fundamental Rights of the European Union and Articles 12 - 15 of the E-Commerce Directive. At the international level, Article 19 of the Universal Declaration of Human Rights and Article 19 of the International Covenant on Civil and Political Rights (ICCPR) are of utmost importance, as well as the 2017 Joint Declaration on “Fake News”, Disinformation and Propaganda that was adopted by the International Specialised Mandates on Freedom of Expression and/or the Media.

Chapter 4 provides an overview of the existing methods and adopted strategies aimed at countering “fake news”, at the European level and in various countries other than the Netherlands. The report distinguishes between preventive or pre-emptive measures; identification and monitoring measures; containing or corrective measures and regulatory and non-regulatory counter-measures. Following this overview, Chapter 5 analyses the effectiveness of these different measures. Importantly, most measures present advantages and disadvantages and, therefore, the focus should be on combining them in an effective way

without infringing individuals' right to freedom of expression. In light of this, preference is given to counter "fake news" on the basis of existing legislation instead of on the basis of new regulatory measures in which such type of content forms the basis for regulation.

Chapter 6 deals with the suitability of measures and initiatives developed elsewhere for replication in the Netherlands. The experiences and controversies of recent regulatory initiatives in Germany, Italy and France provide relevant cautionary lessons for the Netherlands. Continued investment in independent, quality journalism and public service media and content offer longer-term solutions for countering and pre-empting "fake news" and other forms of online disinformation. Having regard to recent initiatives in the Netherlands, it is concluded that media literacy is already very much present; that fact-checking measures and the contextualisation of "fake news" could be more encouraged. Furthermore, non-regulatory measures, such as the use of credibility scores, offer an alternative to regulatory measures.

1. Introduction

Background

“Fake news” has been commanding much attention in the media and in political debate. Towards the end of 2017, the Minister for Primary and Secondary Education and the Media, Mr. Slob, acknowledged that he shares concerns about the dissemination of fake news, “given that various parties have noted in the past years that this is a potential threat for Dutch society”. The Minister also stated his intention to investigate in the coming period the extent to which fake news exists in the Netherlands as well as the effectiveness of relevant initiatives and to discuss the findings with the sector.

It is against this backdrop that the present research was commissioned. At the end of February 2018, the Ministry for Education, Culture and Science invited the Institute for Information Law (IViR) to submit a proposal for an ‘Inventory of methods to counter fake news’. IViR’s proposal was subsequently approved. The parameters of the research, as formulated in the Ministry’s invitation to tender, are set out below. The excerpts have been translated into English by the research team.

~ ~ ~

Aim

Although the phenomenon fake news is much older, the societal discussion about the dissemination of fake news is complex and recent in nature. In the digital society, there are various stakeholders in society and the debate touches on various research disciplines. The aim of this assignment is to present an academic state-of-the-art on the topic, “methods to counter the spread of fake news”. The research is intended to make a contribution to the societal debate on the fake news phenomenon.

Methodology

A literature study based on academic publications, supplemented as relevant with a study of policy, publications of monitoring/regulatory bodies or non-profit research entities.

A Definition

Give an overview of current definitions of the topic with a reasoned preference for a definition or description of the phenomenon fake news.

B Research questions

The study will answer the following questions.

Question 1

Which methods/measures are there to counter fake news at the international level?

Regard will be had here for the actions of different actors, online platforms through which lots of news are spread and viewed, news media, NGOs and governments. If relevant, a reasoned selection of countries may be given. See also Question 3 in this connection.

Possible relevant methods/measures:

- changes of/to online platforms or products where news is presented
- media literacy / information-raising campaigns
- action taken by journalism organisations
- (self-)regulation

Question 2

What is known about the effectiveness of these methods/measures?

The societal discussion about the dissemination of fake news is a recent development which involves many different aspects and which can be countered in different ways. The answer to this question will consider the rationale behind the methods/measures used. What are their desired outcomes and effects? Which actors and instruments are involved? The description should be inventory in nature and does not have to provide definitive conclusions or policy recommendations.

Question 3

How can the research findings be qualified in terms of their applicability to and relevance for the Dutch context?

The answer to this question will consider the legal and social differences with the Netherlands that are relevant for the effectiveness of the described methods/measures.

~ ~ ~

In addition to the above focuses, which were set out by the Ministry, this report will also provide a short overview and analysis of relevant European and international regulation, case-law and policy instruments. This is intended to provide a law and policy framework. It is important to trace this framework at the outset and to take it into account when evaluating the effectiveness of the selected methods/measures to counter fake news. Such methods/measures must also comply with European and international legal standards. The answer to Question 1 will therefore consider relevant instruments and initiatives by European and international organisations, such as the Council of Europe and the European Union.

It is important to emphasize at the outset that this report is an inventory of methods/measures against fake news that have mainly been developed at the European level and/or in other countries. It is not an inventory and analysis of such methods/measures in the Netherlands – that would go beyond the present research assignment. Nevertheless, it is necessary, particularly in Chapter 6 – Applicability in the Netherlands, to briefly discuss a selection of Dutch strategies and methods/measures to counter “fake news”.

2. Definitions

The term, “fake news”, is an amalgamation of different types of expression that are false. Not only is there no authoritative or widely-used definition of the term in academic literature or in law- and policy-making at the international, European or national levels, the term tends to be given different meanings by different commentators. The vagueness of the term, coupled with the divergent meanings ascribed to it, makes it difficult to use it as a stable reference point for research or policy-oriented activities.

Based on an examination of 34 academic articles that used the term “fake news” between 2003 and (early) 2017, Tandoc *et al.* devised a “typology of scholarly definitions of the term”: news satire, news parody, news fabrication, photo manipulation, advertising and public relations, propaganda.¹ While interesting, the typology is of limited use and relevance for present purposes, as only one of the articles surveyed appears to relate to a European country (Italy) and the methodology seems limited (Google Scholar and a library database, with no evidence of having searched in numerous languages). The study does, however, contain a useful focus on the distinctiveness of news, alongside other forms of disinformation (see further, below).

The infographic, ‘Beyond ‘Fake News’ - 10 Types of Misleading News’, produced by the European Association for Viewers Interests (EAVI), is a useful attempt to break down “fake news” into its component parts.² It has deservedly received wide attention and positive acclaim. The infographic is detailed, layered and thoughtful. First, it prises open the term by identifying ten categories of misleading news. It then explores the (likely) motivation behind each category and their (likely) impact. The ten types of misleading news are: propaganda, clickbait, sponsored content, satire and hoax, error, partisan content, conspiracy theory, pseudoscience, misinformation and bogus content. The different types of motivation are listed as: money, politics/power, humour/fun, passion and [the aim to] (mis)inform. Impact is measured on a scale of neutral-low-medium-high.

¹ Edson C. Tandoc Jr., Zheng Wei Lim & Richard Ling (2017): Defining “Fake News”, *Digital Journalism*.

² Available at: <https://eavi.eu/beyond-fake-news-10-types-misleading-info/>. EAVI is an international non-profit organisation.

Other studies also favour disaggregation of the term. In her study, *Fake News. It's Complicated*, Claire Wardle identifies seven types of mis- or disinformation in her breakdown of the term, “fake news”. They are: satire or parody, misleading content, imposter content, fabricated content, false connection, false context and manipulated content.

In their *Information Disorder* study for the Council of Europe, Wardle and Derakhshan refrain from using the term, but seek to position it in a wider frame of mis-, dis- and mal-information. They explain that “it’s important to distinguish messages that are true from those that are false, and messages that are created, produced or distributed by “agents” who intend to do harm from those that are not:

- Dis-information. Information that is false and deliberately created to harm a person, social group, organization or country.
- Mis-information. Information that is false, but not created with the intention of causing harm.
- Mal-information. Information that is based on reality, used to inflict harm on a person, organization or country.”³

In its *Position Paper: “Fake news” and the information disorder*, the European Broadcasting Union (EBU) endorses this approach. It recommends limiting the notion of “fake news” and states that it generally prefers “the term ‘online disinformation’, which is more precise and may serve to designate digital content which contains inaccurate information and which is presented, promoted or disseminated by one or more actors in the chain with the intention to cause harm or make a profit”.⁴ It sees disinformation as “a core problem”, but “only one component of the broader information disorder”.⁵

Other publications use similar continuums of terminology for “these problematic forms of information – misinformation, disinformation, propaganda, gaslighting, and the like”, while emphasizing the need to distinguish between them and to take contextual factors into

³ Claire Wardle and Hossein Derakhshan, ‘Information Disorder: Toward an interdisciplinary framework for research and policy making’, Council of Europe report DGI(2017)09, p. 20 (hereafter, ‘Information Disorder’ study).

⁴ EBU, Position Paper: “Fake news” and the information disorder’, 18 April 2018, p. 6.

⁵ *Ibid.*

account.⁶ It is also pertinent to unpack the different types of mis- and dis-information that can be used for different strategic communication purposes by different actors.⁷

In its Report, *A multi-dimensional approach to disinformation*, the independent High Level Expert Group (HLEG) on fake news and online disinformation observes that debates about “fake news” “encompass a spectrum of information types”. They include: “relatively low-risk forms such as honest mistakes made by reporters, partisan political discourse, and the use of click bait headlines, to high-risk forms such as for instance foreign states or domestic groups that would try to undermine the political process in European Member States and the European Union, through the use of various forms of malicious fabrications, infiltration of grassroots groups, and automated amplification techniques”.⁸ As it “deliberately avoid[s]” the term “fake news” in its report, the HLEG then defines disinformation – which it considers to be a wider term and a more appropriate focus - as “false, inaccurate, or misleading information designed, presented and promoted to intentionally cause public harm or for profit”.⁹

This study uses the term “fake news” only in the interests of convenience and consistency with the assigned research questions. Beyond the confines of the present research, it recommends taking deliberate distance from the term, “fake news”. It does so for conceptual and consequentialist reasons.¹⁰ Conceptually, the term is too vague to be analytically useful and the superficial neatness of the term suggests a coherence that is not there. It is also ill-suited for dealing with the multi-dimensional complexities involved, as has been pointed out by Wardle and Derakhshan and by the HLEG. It is more purposeful to disaggregate the specific types of expression lurking behind the term and try to determine which of them are actually problematic from a democratic perspective and then calibrate appropriate responses to the problematic ones.

⁶ Caroline Jack, *Lexicon of Lies: Terms for Problematic Information*, Data & Society Research Institute, 2017, p. 14.

⁷ Monroe E. Price, *Free Expression, Globalism and the New Strategic Communication* (New York, Cambridge University Press, 2015).

⁸ HLEG, ‘A multi-dimensional approach to disinformation’, Final Report, March 2018, p. 10 (hereafter, HLEG Final Report).

⁹ *Ibid.*

¹⁰ The following two paragraphs draw on Tarlach McGonagle, “‘Fake news’: False fears or real concerns?”, 35 *Netherlands Quarterly of Human Rights* (no. 4, December 2017), 203-209.

From a consequentialist point of view, the term “fake news” can be seen as a term “from the traditional lexicon of autocracy”.¹¹ Politicians, government ministers – and even prime ministers and presidents – sometimes use the term to undermine the reputation and credibility of individual journalists, individual media organisations and the media generally. President Trump’s ongoing vilification of the “fake news media” – a net which he has cast very widely to include reputable media organisations – is a case in point.¹² Such verbal violence against journalists and the media is a very dangerous development for any democratic society. Recognising this, the Council of Europe’s Committee of Ministers has stated forcefully that: “State officials and public figures should not undermine or attack the integrity of journalists and other media actors, for example on the basis of their gender or ethnic identity, or by accusing them of disseminating propaganda, and thereby jeopardise their safety”.¹³ In the same vein, the UN High Commissioner for Human Rights, Zeid Ra’ad Al Hussein, has questioned whether President Trump’s targeting of individual journalists with accusations of propagating “fake news” “[is not] an incitement for others to attack journalists?” He has also described President Trump’s “demonization of the press” (in the US) as “poisonous” because of the consequences it has elsewhere.¹⁴ In sum, the political misappropriation of the term “fake news” is another reason to abandon it, as Wardle and Derakhshan¹⁵ and the HLEG¹⁶ have also pointed out.

Nevertheless, this study understands “fake news” as information that has been deliberately fabricated and disseminated with the intention to deceive and mislead others into believing falsehoods or doubting verifiable facts;¹⁷ it is disinformation that is presented as, or is likely to be perceived as, news. Unlike many other types of information, news plays a particular role in democratic societies: it is a key source of accurate information about political and societal affairs, which informs public opinion-making and deliberative processes. If news is “fake”, it mis-informs the public and democratic debate is polluted at source.

¹¹ Matthew D’Ancona, *Post Truth: The New War on Truth and How to Fight Back* (Penguin, London, 2017), p. 58.

¹² Joel Simon, ‘Trump Is Damaging Press Freedom in the U.S. and Abroad’, *The New York Times*, 25 February 2017, available at: <https://www.nytimes.com/2017/02/25/opinion/trump-is-damaging-press-freedom-in-the-us-and-abroad.html?mcubz=3>

¹³ Guidelines in Appendix to Recommendation CM/Rec(2016)4 of the Committee of Ministers to member States on the protection of journalism and safety of journalists and other media actors, 13 April 2016, para. 15.

¹⁴ UN High Commissioner for Human Rights, Statement, 30 August 2017, available at: <http://www.unmultimedia.org/avlibrary/asset/1958/1958570/>.

¹⁵ ‘Information Disorder’ study, p. 16.

¹⁶ HLEG Final Report, p. 10.

¹⁷ The first prong of this definition has been borrowed, almost verbatim, from the Ethical Journalism Network: <http://ethicaljournalismnetwork.org/tag/fake-news>.

News as a distinctive genre of information, a *primus inter pares* in democratic society, in view of its role to provide members of society with accurate information about public affairs. Michael Schudson, a leading academic expert on the value of news in democratic society, has identified seven goals a media system dedicated to democracy might aspire to. According to him, the (news) media should:

1. Provide citizens fair and full information so that they can make sound decisions as citizens.
2. Provide coherent frameworks to help citizens comprehend the complex political universe.
3. Serve as common carriers of the perspectives of the varied groups in society.
4. Provide the quantity and quality of news that people want.
5. Represent the public and speak for and to the public interest in order to hold government accountable.
6. Evoke empathy and provide deep understanding so that citizens at large can appreciate the situation of other human beings in the world and so elites can come to know and understand the situation of other human beings, notably nonelites, and learn compassion for them.
7. Provide a forum for dialogue among citizens that not only informs democratic decision making but is, as a process, an element in it.¹⁸

It is becoming increasingly difficult for news media to realise these objectives in the present multi-media environment. Context is key; news as a genre is not just a string of separate items, which can be packaged and distributed swiftly and often very widely as stand-alone items. There is increased movement towards this kind of “distributed discovery”, in which social media, search engines and aggregators, more and more determine the terms and nature of the distribution of news.¹⁹ In this changed multi-media environment, one must also be aware that: “A diverse combination of actors including **trolls, bots, fake-news websites, conspiracy theorists, politicians, highly partisan media outlets, the mainstream media,**

¹⁸ Abridged list from Michael Schudson, *The Power of News* (Cambridge, Massachusetts and London, England, Harvard University Press, 1995), pp. 28-29.

¹⁹ Rasmus Kleis Nielsen, Alessio Cornia and Antonis Kalogeropoulos, ‘Challenges and opportunities for news media and journalism in an increasingly digital, mobile and social media environment’, Council of Europe report DGI(2016)18, Strasbourg, October 2016, p. 12.

and foreign governments are all playing overlapping—and sometimes competing—roles in producing and amplifying disinformation in the modern media ecosystem”.²⁰ And in “the subsequent cacophony, the flow of information is increasingly dominated by peer-to-peer interaction rather than the imprimatur of the traditional press”.²¹ A further complicating factor is the growth of computational propaganda, described/defined as “the use of algorithms, automation, and human curation to purposefully distribute misleading information over social media networks”.²²

Besides the different actors and different technological means used to create, amplify and disseminate “fake news” and other forms of disinformation, it is also important to differentiate between the different motivations behind such types of disinformation. They often involve conspiracy theories, for instance about climate change, immigration, vaccinations, historical events and political affairs. They often involve propaganda designed to disrupt elections and referenda in the countries where they are created and/or in foreign countries. Clickbait for financial gain is another frequent motivation for these types of disinformation. Clickbait is understood as online content that primarily aims to lure readers to a particular website, typically through clicking on misleading, sensationalist or provocative titles. So-called “fake news” farms in Macedonia were reported to have generated such content in a lucrative business model during the US Presidential elections.²³ Political and financial motivation can also dovetail.

In recent years, social media have seen a rapid growth in usage for news consumption. The first main finding of the Reuters Institute’s Digital News Report 2016 was: “across our entire sample, half (51%) say they use social media as a source of news each week”. The 2017 Report finds that this growth is flattening in some markets, such as Sweden and Germany. A reason for this could be the increased popularity of messaging apps such as WhatsApp for news. WhatsApp’s popularity as news source is mainly outside of Europe (Latin America and South-East Asia) but in Spain, use of the messaging app for news has increased from 24% to

²⁰ (emphasis per original) Joshua A. Tucker *et al.*, ‘Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature’, Hewlett Foundation, 19 March 2018, p. 22.

²¹ Matthew D’Ancona, *Post Truth: The New War on Truth and How to Fight Back*, *op. cit.*, p. 53.

²² Samuel C. Woolley and Philip N. Howard, ‘Computational Propaganda Worldwide’, Oxford, Oxford Internet Institute, University of Oxford, Executive summary, p. 6. See also: Samantha Bradshaw and Philip N. Howard, (2017) ‘Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation’, Oxford Computational Propaganda Project, Working paper no. 2017.12, Oxford, 17 July 2017.

²³ See, for example, Emma Jane Kirby, ‘The city getting rich from fake news’, BBC News, 6 December 2016, available at: <http://www.bbc.com/news/magazine-38168281>.

32%. Facebook's global popularity is however unrivalled, with both the social network and its accompanying messaging app being used as news media. 80% of the respondents use a Facebook product weekly for any purpose.

Participation in online news varies substantially by country but in general, sharing or commenting on news has been stationary in most countries over the past year. In southern European countries such as Portugal and Italy, a relatively high percentage of the population shares news weekly: 51% and 47% respectively. This includes both sharing on social media and on websites of news organisations. In northern European countries, this percentage is generally lower: 18% in Germany and 22% in the Netherlands. The main reasons for not sharing or commenting are a lack of interest and a preference to discuss news stories face-to-face.

It has been noted that the lack of definitions can be a hindrance to regulation, but whether new or specific regulation is required is another question entirely. Chapter 2 sets out the existing European law and policy framework governing freedom of expression, with emphasis on "fake news" and disinformation.

3. European (legal and policy) framework

This chapter briefly sets out the European law and policy framework on freedom of expression, media freedom and pluralism, which shapes the regulatory context in which measures against “fake news” must be considered.

Council of Europe standards

Article 10 of the European Convention on Human Rights (ECHR or the Convention) is the centrepiece of the Council of Europe’s system for the protection of the right to freedom of expression. It reads:

1. Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers. This Article shall not prevent States from requiring the licensing of broadcasting, television or cinema enterprises.
2. The exercise of these freedoms, since it carries with it duties and responsibilities, may be subject to such formalities, conditions, restrictions or penalties as are prescribed by law and are necessary in a democratic society, in the interests of national security, territorial integrity or public safety, for the prevention of disorder or crime, for the protection of health or morals, for the protection of the reputation or rights of others, for preventing the disclosure of information received in confidence, or for maintaining the authority and impartiality of the judiciary.

Article 10 § 1 sets out the right to freedom of expression, which comprises the freedom to hold opinions, the freedom to receive information and ideas and the freedom to impart information and ideas. Under Article 10 § 1, it is possible for states to regulate audiovisual media by means of licensing schemes.

Article 10 § 2 delineates the scope of the core right set out in the preceding paragraph. It does so by enumerating a number of grounds, based on which the right may legitimately be restricted, provided that the restrictions are “prescribed by law” and are “necessary in a democratic society”. It justifies this approach by linking the permissibility of restrictions on

the right to freedom of expression to the existence of “duties and responsibilities” that govern its exercise. The scope of those duties and responsibilities varies, depending on the “situation” of the person exercising the right to freedom of expression and on the “technical means” used.²⁴ The Court has tended to explore the nature and scope of relevant duties and responsibilities not through broad principles, but on a case-by-case basis. It tends to distinguish among different professional occupations, such as journalism, politics, education and military service.

Article 10, as interpreted by the Court, provides strong protection to the right to freedom of expression. The Court consistently describes the right as “one of the essential foundations of a democratic society and one of the basic conditions for its progress and for each individual’s self-fulfilment”.²⁵ As the Court affirmed in its seminal judgment in *Handyside v. the United Kingdom*, freedom of expression “is applicable not only to ‘information’ or ‘ideas’ that are favourably received or regarded as inoffensive or as a matter of indifference, but also to those that offend, shock or disturb the State or any sector of the population. Such are the demands of that pluralism, tolerance and broadmindedness without which there is no ‘democratic society’” (§ 49). This principle creates the necessary space for robust, pluralistic public debate in democratic society.

The Court has developed a standard test to determine whether Article 10 of the Convention has been violated. Put simply, whenever it has been established that there has been an interference with the right to freedom of expression, that interference must first of all be prescribed by law (i.e., it must be adequately accessible and reasonably foreseeable in its consequences). Second, it must pursue a legitimate aim, that is correspond to one of the aims set out in Article 10 § 2. Third, the interference must be necessary in a democratic society, which means it must correspond to a “pressing social need”, and be proportionate to the legitimate aim(s) pursued.

Under the margin of appreciation doctrine, which takes account of how the Convention is interpreted at national level, states are given a certain amount of discretion in how they

²⁴ *Fressoz and Roire v. France* [GC], no. 29183/95, § 52, ECHR 1999-I.

²⁵ See, for an early authority, *Lingens v. Austria*, 8 July 1986, § 41, Series A no. 103, paraphrasing *Handyside v. the United Kingdom*, 7 December 1976, § 49, Series A, no. 24.

regulate expression.²⁶ The extent of this discretion, which is subject to supervision by the Court, varies depending on the nature of the expression in question. Whereas states only have a narrow margin of appreciation in respect of political expression, they enjoy a wider margin of appreciation in respect of public morals, decency and religion. Lastly, it is also worth noting that the “national margin of appreciation is circumscribed by the interest of democratic society in enabling the press to exercise its vital role of ‘public watchdog’”,²⁷ a consideration that “weigh[s] heavily” in the balancing exercise.²⁸

Besides the margin of appreciation doctrine, three other interpretive principles espoused by the Court are of particular relevance for the right to freedom of expression. These are the practical and effective doctrine, the living instrument doctrine and the positive obligations doctrine. According to the practical and effective doctrine, all rights guaranteed by the Convention must be “practical and effective” and not merely “theoretical or illusory”.²⁹ Under the “living instrument” doctrine, the Convention is regarded as a “living instrument” that “must be interpreted in the light of present-day conditions”.³⁰ This doctrine seeks to ensure that the Convention evolves with the times and does not become static or outdated. The positive obligations doctrine implies that it is not always enough for the state to simply refrain from interfering with individuals’ human rights: positive or affirmative action will often be required as well. Thus, notwithstanding the tendency to formulate states’ obligations in negative terms, in order to ensure that the rights enshrined in the Convention are practical and effective, states may have to take positive measures, “even in the sphere of the relations of individuals between themselves”.³¹ Relevant positive obligations for present purposes include guaranteeing pluralism in the media sector³² and fostering a favourable environment for for public debate.³³

The right to freedom of expression, as guaranteed by Article 10 ECHR, is not limited to protection for truthful information. The Court has held in this respect that Article 10 “does not prohibit discussion or dissemination of information received even if it is strongly

²⁶ Initially developed in the Court’s case law (see, in particular: *Handyside v. the United Kingdom*, *op. cit.*, §§ 47 to 50).

²⁷ *Chauvy and Others v. France*, no. 64915/01, § 67, ECHR 2004-VI.

²⁸ *Colombani and Others v. France*, no. 51279/99, § 57, ECHR 2002-V.

²⁹ *Airey v. Ireland*, 9 October 1979, § 24, Series A, no. 32.

³⁰ *Tyrer v. the United Kingdom*, 25 April 1978, § 31, Series A no. 26.

³¹ *X and Y v. the Netherlands*, 26 March 1985, § 23, Series A no. 91.

³² *Informationsverein Lentia and Others v. Austria*, 24 November 1993, § 38, Series A no. 276.

³³ *Dink v. Turkey*, nos. 2668/07, 6102/08, 30079/08, 7072/09 and 7124/09, § 137, 14 September 2010.

suspected that this information might not be truthful. To suggest otherwise would deprive persons of the right to express their views and opinions about statements made in the mass media and would thus place an unreasonable restriction on the freedom of expression set forth in Article 10 of the Convention”.³⁴ The Court has, however, consistently stressed that facts and opinions or value judgments are not the same. The existence of facts can be demonstrated, but it is not possible to prove the truth of opinions or value judgments.³⁵ A requirement to prove the truth of a value judgment infringes the right to freedom of opinion. A value judgment should, however, have adequate factual basis, as even a value judgment without any factual basis to support it may be excessive.³⁶

In its case-law, the Court has developed a number of principles to enable journalists and the media to fulfil the important tasks ascribed to them in democratic society, namely to:

- disseminate information and ideas widely and thereby inform and influence public opinion-making;
- act as public watchdogs holding authorities to account, and
- create shared forums and channels through which public debate takes place.

The principles that the Court has developed seek to safeguard the editorial freedom and operational autonomy of journalists and the media.³⁷

In response to technology-driven developments in how we communicate with each other, the Court has come to recognise that an increasing range of actors – and not just journalists and media professionals - nowadays contribute to public debate, including whistle-blowers, citizen journalists, bloggers, civil society organisations, academics and individual commentators. It has recognised the valuable contribution that these types of actors can make to public debate. Following the logic of this growth-curve in the Court’s jurisprudence, it is likely that fact-checking organisations will be given similar recognition in the future, insofar as they are another form of what is sometimes referred to as “accountability journalism”.³⁸

The Court has also recognised how important the Internet has become for information and communication, which requires clear regulatory frameworks. But it has also observed that

³⁴ *Salov v. Ukraine*, no. 65518/01, § 113, ECHR 2005-VIII.

³⁵ *Lingens v. Austria*, 8 July 1986, § 46, Series A no. 103.

³⁶ *Dichand and Others v. Austria*, no. 29271/95, §§ 42 and 43, 26 February 2002.

³⁷ See, amongst many other judgments, *Jersild v. Denmark*, 23 September 1994, Series A no. 298.

³⁸ C.W. Anderson, Leonard Downie Jr. and Michael Schudson, *The News Media: What Everyone Needs to Know* (New York, Oxford University Press, 2016), p. 112.

much online expression is vulgar and offensive. It recognises that Internet intermediaries wield great control over the flow and content of online expression, describing them as “protagonists of the free electronic media”,³⁹ and drawing attention to the duties and responsibilities that govern their gate-keeping and other activities.

In short, the Court is – after an initial delay – increasingly addressing the complexities and fast-evolving nature of the contemporary multi-actor and multi-level media ecosystem. Monroe Price captures these complexities and transformations very well when he draws attention to the “[i]mportant shifts [that] take place between self-generated or civil-society generated contributions and the directed, overt and disciplined imposition of information flows from highly organized strategic communicators”.⁴⁰ This is the contemporary communications context in which “fake news” and online disinformation have become prevalent and the context for which the Court’s long-standing and emerging principles have to be (re-)applied.

Finally, in respect of the principles developed by the Court in its interpretation of Article 10 ECHR, it is relevant to refer to elections as there are particular concerns about the impact of “fake news” and online disinformation on election processes. The Court has held that “[f]ree elections and freedom of expression, particularly freedom of political debate, together form the bedrock of any democratic system”.⁴¹ It continued: “[t]he two rights are inter-related and operate to reinforce each other... For this reason, it is particularly important in the period preceding an election that opinions and information of all kinds are permitted to circulate freely”.⁴² At the same time, the Court has stressed “it remains the case that both during and outwith an election period, the print media’s activity is subject to the requirement to act in good faith in order to provide accurate and reliable information in accordance with the ethics of journalism [...] and considerations relating to certain boundaries, particularly as regards the reputation and rights of others and the need to prevent the disclosure of confidential information”.⁴³

³⁹ *Magyar Tartalomszolgáltatók Egyesülete and Index.hu Zrt v. Hungary*, no. 22947/13, § 88, 2 February 2016.

⁴⁰ Monroe E. Price, *Free Expression, Globalism and the New Strategic Communication*, *op. cit.*, p. 27.

⁴¹ *Bowman v. the United Kingdom*, 19 February 1998, § 42, Reports of Judgments and Decisions 1998-I.

⁴² *Ibid.*

⁴³ *Orlovskaya Iskra v. Russia*, no. 42911/08, § 131, 21 February 2017.

Besides the ECHR and the case-law of the Court, the Council of Europe also has an extensive body of political standards that help to shape the organisation's approach to "fake news" and online disinformation. The Secretary General of the Council of Europe recently drew considerable attention to these issues in the broader context of human rights and the rise of populism in Europe.⁴⁴ The bulk of these political standards has been developed by the organisation's Committee of Ministers. Most recently, the Committee of Ministers has addressed Recommendations to the organisation's 47 member States on: media pluralism and transparency of media ownership,⁴⁵ and the roles and responsibilities of internet intermediaries.⁴⁶ Previous Recommendations and Declarations have addressed such topics as: freedom of political debate in the media;⁴⁷ a new notion of media;⁴⁸ social networking services and human rights,⁴⁹ and search engines and human rights.⁵⁰ Ongoing standard-setting work includes a focus on how to promote a favourable environment for quality journalism in the digital age.

European Union standards

The Charter of Fundamental Rights of the European Union is the EU's flagship instrument for the protection of human rights. The Charter's provisions "are addressed to the institutions, bodies, offices and agencies of the Union with due regard for the principle of subsidiarity and to the Member States only when they are implementing Union law" (Article 51(1)). The Charter's provisions which "contain principles may be implemented by legislative and executive acts taken by institutions, bodies, offices and agencies of the Union, and by acts of Member States when they are implementing Union law, in the exercise of their respective powers" (Article 52(5)). However, they shall be "judicially cognisable only in the interpretation of such acts and in the ruling on their legality" (*ibid.*).

⁴⁴ Secretary General of the Council of Europe, 'Populism – How strong are Europe's checks and balances?', 4th annual report on the state of democracy, human rights and the rule of law in Europe, 2017, Chapter 2 – Freedom of Expression.

⁴⁵ Council of Europe, Recommendation CM/Rec(2018)1 of the Committee of Ministers to member States on media pluralism and transparency of media ownership, 7 March 2018.

⁴⁶ Council of Europe, Recommendation CM/Rec(2018)2 of the Committee of Ministers to member States on the roles and responsibilities of internet intermediaries, 7 March 2018.

⁴⁷ Council of Europe, Committee of Ministers' Declaration on freedom of political debate in the media, 12 February 2004.

⁴⁸ Council of Europe, Recommendation CM/Rec(2011)7 of the Committee of Ministers to member states on a new notion of media, 21 September 2011.

⁴⁹ Council of Europe, Recommendation CM/Rec(2012)4 of the Committee of Ministers to member States on the protection of human rights with regard to social networking services, 4 April 2012.

⁵⁰ Council of Europe, Recommendation CM/Rec(2012)3 of the Committee of Ministers to member States on the protection of human rights with regard to search engines, 4 April 2012.

It is important to ensure that the human rights standards elaborated by the Council of Europe and the EU are broadly consistent or equivalent. Thus, the Charter provides that insofar as the Charter contains rights that correspond to those safeguarded by the European Convention on Human Rights, “the meaning and scope of those rights shall be the same as those laid down by” the Convention (Article 52(3)). Article 11 of the Charter – which focuses on freedom of expression, as well as media freedom and pluralism - should therefore be interpreted consistently with Article 10 of the Convention and relevant case-law of the European Court of Human Rights. The text of Article 11 of the Charter is in any case modelled on Article 10 of the Convention, but more succinctly formulated. All of this means that the principles from relevant case-law of the European Court of Human Rights (set out above) ought to govern the interpretation of Article 11 of the Charter by the Court of Justice of the European Union (CJEU).

However, notwithstanding the broad consistency between the Convention and the Charter, the latter explicitly recognises a number of rights that are not included in the Convention or the Protocols thereto. It refers to “fundamental” rights instead of “human” rights, which explains the inclusion of fundamental (economic) rights that are central to the EU’s core values as an economic community espousing an internal market. The freedom to conduct a business, safeguarded by Article 16 of the Charter, is a good example.

Besides the Charter, other regulatory instruments are also relevant. The EU’s E-Commerce Directive establishes exemptions from liability for Internet service providers when they meet certain conditions.⁵¹ The exemptions are set out in Articles 12-14 of the Directive and they can be availed of by service providers acting as a ‘mere conduit’ for information, or those which provide ‘caching’ or ‘hosting’ services. This means that hosting service providers would ordinarily benefit from an exemption for liability for illegal content, as long as they maintain a neutral or passive stance towards that content. A service provider that hosts third-party content may avail of this exemption on condition that it does not have “actual knowledge of illegal activity or information and, as regards claims for damages, is not aware of facts or circumstances from which the illegal activity or information is apparent” and that

⁵¹ Recital 42, Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market, [2000] OJ L 178/1 (Directive on Electronic Commerce).

“upon obtaining such knowledge or awareness, acts expeditiously to remove or to disable access to the information”.⁵² However, “the removal or disabling of access has to be undertaken in the observance of the principle of freedom of expression and of procedures established for this purpose at national level”.⁵³ Pursuant to Article 15 of the Directive, EU Member States are not allowed to impose a general obligation on providers to “monitor the information which they transmit or store, nor a general obligation actively to seek facts or circumstances indicating illegal activity”. The type of surveillance that such a general monitoring obligation would entail would have a chilling effect on the freedom of expression of users of the service.

Although merely of advisory value, the HLEG’s recommendations to the European Commission, are likely to prove influential. Its recommendations are based on five pillars, designed to:

- “1. enhance **transparency** of online news, involving an adequate and privacy-compliant sharing of data about the systems that enable their circulation online;
2. promote **media and information literacy** to counter disinformation and help users navigate the digital media environment;
3. develop tools for **empowering users and journalists** to tackle disinformation and foster a positive engagement with fast-evolving information technologies;
4. safeguard **the diversity and sustainability of the European news media ecosystem**, and
5. promote **continued research** on the impact of disinformation in Europe to evaluate the measures taken by different actors and constantly adjust the necessary responses.”⁵⁴

The European Commission is expected to issue a new Communication on “fake news” and online disinformation on 25 April. Commission Communications are not legally binding on Member States; they are a means for the Commission to develop its position on particular – usually topical – issues. The forthcoming Communication is expected to draw on the HLEG’s final report and (to a lesser extent) the outcome of the public consultation on the issue. It is also expected to set out a possible European-level approach.⁵⁵

⁵² Article 14, *ibid.*

⁵³ Recital 46, *ibid.*

⁵⁴ [Bold per original] HLEG Final Report, pp. 5-6.

⁵⁵ Please note that the European Commission’s Communication, ‘Tackling online disinformation: a European Approach’, COM(2018) 236 (final) was issued on 26 April 2018 – after the present study was completed. For

International standards

Within the international human rights system, the central provisions for the protection of freedom of expression are Article 19 of the Universal Declaration of Human Rights and Article 19 (*juncto* Article 20) of the International Covenant on Civil and Political Rights (ICCPR). As these provisions are broadly similar to Article 10 ECHR, they will not be explored in depth here. It should, however, be noted that the Human Rights Committee – the oversight body of the ICCPR, has in the past expressed its deep concern “at the prosecution and punishment of journalists for the crime of publication of false news merely on the ground, without more, that the news was false, in clear violation of article 19 of the Covenant”.⁵⁶

Also noteworthy is the Joint Declaration on “Fake News”, Disinformation and Propaganda that was adopted by the International Specialised Mandates on Freedom of Expression and/or the Media in March 2017.⁵⁷ The Joint Declaration is part of an annual series; each Joint Declaration examines different topics relating to freedom of expression and/or the media. These Joint Declarations are not legally binding, but owing to their collective endorsement by the specialised mandates, they are regarded as very persuasive interpretations of existing international human rights law on the topics they address.⁵⁸ The 2017 Joint Declaration is the most explicit and detailed international text addressing “fake news” in recent years. It puts much store by the need for “fake news” to be dealt with in the context of an enabling environment for free expression. It also includes pertinent focuses on the roles and responsibilities of States, Internet intermediaries, journalists and media outlets and other stakeholders in responding to “fake news”. It is noteworthy that the specialised mandates use the term within scare quotes, indicating a certain reluctance to embrace the term, perhaps, or

the readers’ convenience, the European Commission’s press release announcing the Communication has been appended to the present study.

⁵⁶ UN Human Rights Committee: Concluding Observations: Cameroon, 4 November 1999, CCPR/C/79/Add.116, para. 24.

⁵⁷ The United Nations (UN) Special Rapporteur on Freedom of Opinion and Expression, the Organization for Security and Co-operation in Europe (OSCE) Representative on Freedom of the Media, the Organization of American States (OAS) Special Rapporteur on Freedom of Expression and the African Commission on Human and Peoples’ Rights (ACHPR) Special Rapporteur on Freedom of Expression and Access to Information, Joint Declaration on “Fake News”, Disinformation and Propaganda, 3 March 2017.

⁵⁸ For analysis, see: Toby Mendel, ‘The UN Special Rapporteur on freedom of opinion and expression: progressive development of international standards relating to freedom of expression’, in T. McGonagle and Y. Donders, Eds., *The United Nations and Freedom of Expression and Information: Critical Perspectives* (Cambridge, Cambridge University Press, 2015), pp. 235-268, esp. at 251-257.

at least to underline its contested nature. It is also telling that the Joint Declaration does not address “fake news” exclusively: it positions the term on a continuum with other types of disinformation.

4. Methods/measures against “fake news”

Given the multi-faceted nature of “fake news” and online disinformation, it is no surprise that a range of strategies and measures already exist to tackle such kinds of disinformation. This chapter seeks to explore a representative selection of relevant strategies and measures that are up and running at the European level and in various countries (other than the Netherlands). The exploration is intended as an overview, not as a comprehensive or exhaustive list. Readers are encouraged to read this section in the light of the Information Disorder study for the Council of Europe and the Multi-dimensional approach report by the HLEG, as both set out extensive lists, which the present research has partly drawn on.

An overview of the surveyed strategies and measures could be grouped in different ways, such as by type, by objective or by actor. To organise the various measures “by type” would be the most straightforward categorisation and lead to a simple inventory. However, there is added analytical value in grouping them by objective or by actor, even if some measures may have more than one objective or different actors carry out various types of measures. Classification of measures as short-, medium- and long-term can also be analytically useful.

The approach taken below is to group the strategies and measures according to the following objectives:

1. Preventive or pre-emptive measures;
2. Identification and monitoring measures;
3. Containing or corrective measures;
4. Regulatory and non-regulatory counter-measures.

Owing to the complexity of “fake news” and disinformation, both the *Information Disorder* study and the HLEG’s Report, insist on the need for multi-actor approach. In the category, “by actor”, distinctive roles could be identified for:

- State authorities (government (ministries) and other public authorities);

- Mainstream media (with appropriate differentiation between various types, such as public service⁵⁹ and commercial);
- Journalists and editors;
- Platform operators;
- Civil society/NGOs;
- Educational establishments;
- Individual citizens.⁶⁰

Collaboration between different actors can unlock synergies in terms of pooling complementary expertise, accessing new networks and diversifying strategies and approaches.

4.1 Preventive or pre-emptive measures

At the beginning of this section, it should be noted that one of the most consistent recommendations for countering “fake news” and other forms of online disinformation is that States should create a favourable environment for independent, quality media and media content, including through financial investment and other support measures. All such measures should be administered in such a way as to guarantee the editorial freedom and operational autonomy of the media.

4.1.1 Media, information and news literacy

We have come a very long way from the (recent) days when it was widely believed that “the camera never lies”. Nowadays, text can be conjured out of nowhere or generated by bots; photographic images can be photoshopped or otherwise doctored; video footage can be dubbed with false texts or otherwise manipulated, and with merely a webcam and appropriate, readily available software, it appears quite easy to also very convincingly impose facial and head movements on footage of another person, for example a world leader.⁶¹ The

⁵⁹ In its position paper, the EBU explains the role of independent public service media in supporting informed citizenship and fighting disinformation. See generally the EBU Position paper, “Fake news” and the information disorder’, *op. cit.*, and in particular, Part Two.

⁶⁰ For detailed recommendations for different actors, see, for example, ‘Information Disorder’ study, pp. 80-85.

⁶¹ For a short demonstration of how this works in practice, see: Justus Thies, Michael Zollhöfer, Marc Stamminger, Christian Theobalt & Matthias Nießner, ‘Face2Face: Real-time Face Capture and Reenactment of

ease with which false content can be generated or authentic content manipulated to convey false or distorted messages and impressions is very troubling. It calls for greater vigilance than ever on the part of public watchdogs and ordinary individuals.

With the increasing development and use of digital technologies in the media environment, there is a pressing need create and sustain critical-thinking skills for analyzing online news content. Citizens who are media literate can ascertain the reliability of news content and sort fact from opinion in order to make more informed choices about their news consumption. News literacy does not solely refer to news content in the form of texts, but also includes educating people on online persuasion tools and the manipulative powers of (moving) images.

Examples:

Educational projects: In Belgium, the Flemish media literacy centre, *Mediawijs*, focused on educating people about how to engage more critically with “fake news” in 2017.⁶² Moreover, every year the media education project, “news in the class”, is organised. It is a collaboration between Mediawijs, Vlaamse Nieuwsmedia, Press and Media, with support from the Flemish government.⁶³ This project aims to stimulate students to consult and interpret news sources in a critical way.⁶⁴ Consequently, it provides teachers with educational packages, which recently also included one on the countering of “fake news”.⁶⁵ Through such packages, students are challenged and learn how to discern false and misleading news and, through such means, they have been exposed to the ‘filter bubble’ phenomenon.⁶⁶ Similar initiatives have been launched in Wallonia by the Superior Council for Media Education, such as the “two weeks of media education”, the

RGB Videos’ (CVPR 2016 Oral), available at: <http://niessnerlab.org/projects/thies2016face.html>. For additional reporting, see: Enith Vlooswijk, ‘Nieuwe software maakt van politici digitale handpoppen’, *De Volkskrant*, 3 June 2017.

⁶² Mediawijs, ‘Dossier: Nieuws- en Informatiegeletterdheid’, 2017, available at: <https://mediawijs.be/dossiers/dossier-nieuws-en-informatiegeletterdheid>.

⁶³ Nieuws in de Klas, ‘Nieuws in de klas. Jouw afspraak met het nieuws!’, available at: <https://www.nieuwsindeklas.be/contact/>.

⁶⁴ *Ibid.*

⁶⁵ Nieuws in de klas, ‘Fake news? What’s in a name’, 2017, available at: https://www.nieuwsindeklas.be/wp-content/uploads/2016/09/Fake-News_lesmap.pdf; Mediawijs, ‘Fake News? What’s in a name’, available at: https://mediawijs.be/sites/default/files/wysiwyg/fake_news_-_difference_day_1.pdf.

⁶⁶ Vlaamse Nieuwsmedia, ‘Nieuws in de klas: Fake News op Difference Day’, 2017, available at: <https://www.vlaamsenieuwsmedia.be/newsroom/2017/04/nieuws-in-de-klas-fake-news-op-difference-day>.

main topic of which in 2017 was “fake news”.⁶⁷ On 18 October 2017, the Council also organized the “day of media education”, which started with a debate around “fake news”.⁶⁸

- Involving citizens in policy-making: on 31 January 2018, Sven Gatz, the Flemish Minister for Culture, announced that he would organize a ‘burgerkabinet’ (citizen cabinet) focusing on ‘fake news’.⁶⁹ This project aims to engage with citizens on how they inform themselves today and on what could be done better.⁷⁰ It was possible to share ideas through an online platform until 7 April 2018.⁷¹ Afterwards, a “real” discussion platform was scheduled for 21 April 2018 in the Flemish parliament. 150 people were expected to be involved with the purpose of issuing policy recommendations.⁷²
- Libraries and the promotion of critical thinking: the International Federation of Library Associations and Institutions (IFLA) has been playing an important role in countering “fake news”, by promoting critical thinking. More specifically, the organisation has developed an infographic setting out eight simple steps on ‘How to Spot Fake News’.⁷³ The steps have been derived from a 2016 article by FactCheck.org. The steps are: consider the source, check the author, check the date, check your biases, read beyond, supporting sources?, is it a joke? and ask the experts.
- The efforts of public service broadcasters: the European Broadcasting Union (EBU) is developing resources for, and a show-case of, its members’ approaches to “fake news” and online disinformation. An overview is available in its Perfect Storm publication.⁷⁴

4.1.2 Trust-enhancing practices

⁶⁷ Conseil Supérieur de l’éducation aux médias, ‘quinzaine de l’éducation aux médias’, available at: www.csem.be/outils/operations/la_quinzaine_de_leducation_aux_medias.

⁶⁸ *Ibid.*

⁶⁹ Open VLD, ‘Sven Gatz lanceert het 4^e Burgerkabinet’, 31 January 2018, available at: <http://www.openvld.be/?type=nieuws&id=1&pageid=96598>.

⁷⁰ *Ibid.*

⁷¹ See: <http://www.burgerkabinet.be/>.

⁷² *Ibid.*

⁷³ See: <https://www.ifla.org/publications/node/11174>.

⁷⁴ Bill Dunlop, *Perfect storm: the multiple challenges facing public service news, and why tackling them is vital for democracy* (Geneva, European Broadcasting Union, 2017), pp. 13 – 15.

Trust-enhancing practices are aimed at strengthening the public's trust and include measures such as increasing transparency and adhering to high ethical and professional standards, as set out in, Codes of Conduct, for example. Online platforms, news publishers and broadcasters as well as civil society organisations have made use of such practices. The public's lack of trust in traditional and new media organisations and actors⁷⁵ has been a contributing cause to the emergence of “fake news”. Building trust takes time and it is not achieved through quick-fix solutions. Trust thrives on connections and engagement with audiences and readers. Trust can be built from within individual media and news organisations, or it can be built collectively at the sectoral level through effective self-regulatory mechanisms.

Examples:

- The Journalism Trust Initiative (JTI) was launched on 3 April by Reporters Without Borders, Agence France Presse, the European Broadcasting Union and the Global Editors Network. The JTI is designed to “promote journalism by adherence to an agreed set of trust and transparency standards to be developed and implemented”.⁷⁶ In that way it will combat disinformation. The standards will be developed in collaboration with stakeholders in the coming period.
- The Ethical Journalism Network (EJN) is a worldwide alliance of reporters, editors and publishers who are committed to promoting accountable journalism.⁷⁷ Its five key principles are: truth and accuracy, independence, fairness and impartiality, humanity and accountability. The EJN believes that “to enforce these core values, newsrooms and media organizations should adopt a codes [sic] of conduct”.⁷⁸ It has developed a very extensive and well-used database of media codes of ethics and press councils, housed at: <https://accountablejournalism.org/>.

4.1.3 Technological initiatives

⁷⁵ See Nic Newman and Richard Fletcher, ‘Bias, Bullshit and Lies: Audience Perspectives on Low Trust in the Media’, Reuters Institute for the Study of Journalism, Oxford, 2017.

⁷⁶ See: <https://rsf.org/en/news/rsf-and-its-partners-unveil-journalism-trust-initiative-combat-disinformation>.

⁷⁷ See: <https://ethicaljournalismnetwork.org/>.

⁷⁸ See: <https://accountablejournalism.org/>.

Although “fake news” is as old as the hills,⁷⁹ it has recently acquired “an alarming new patina”.⁸⁰ Given the scale, speed and sophistication with which the current wave of “fake news” is unfurling, Internet intermediaries are struggling with the variety of ways “fake news” can manifest itself. This has prompted them to seek to develop technological solutions to pre-empt, counter, or contain “fake news”.

Examples:

- Facebook has developed an ‘Election integrity programme’ to act against hacking and malware, to examine the role of ads and foreign interference and to understand fake accounts.⁸¹
- Facebook has developed the ‘ThreatExchange program’, a social platform for sharing cybersecurity threats and information with tech companies on attempts at interference.⁸²
- Twitter uses API-based tools in order to prevent its users from performing coordinated (automated) actions across multiple accounts. These tools limit the possibility for users to be active on different accounts and thereby to perform simultaneous actions such as ‘Likes, Retweets or Follows’.⁸³
- Google project shield: provides News, Human Rights and Elections Monitoring sites with protection from Distributed Denial of Service (DDoS) attacks.⁸⁴

4.2 Identification and monitoring measures

4.2.1 Fact-checking

Fact-checking measures consist in checking the accuracy of online content that is presented as truth or fact in order to debunk disinformation. This can either be done internally by technology companies, by independent external fact-checking organisations, or through collaborations between them and/or other actors. Different actors, such as online platforms,

⁷⁹ See, for example, Adrian Chen, ‘The Fake-News Fallacy’, *The New Yorker*, 28 August 2017 and David Uberti, ‘The real history of fake news’, *Columbia Journalism Review*, 15 December 2016.

⁸⁰ Nicole A. Cooke, ‘Posttruth, Truthiness, and Alternative Facts: Information Behavior and Critical Information Consumption for a New Age’, *The Library Quarterly* 87, no. 3 (July 2017): 211-221, p. 211.

⁸¹ HLEG Final Report, p. 15, footnote 10.

⁸² *Ibid.*

⁸³ *Ibid.*

⁸⁴ *Ibid.*

news media publishers and broadcasters, and news consumers all have had recourse to such practice.⁸⁵

Examples:

- International Fact-Checking Network (IFCN) Code of Principles: This Code of Principles was launched on 15 September 2016 and currently counts forty-eight verified signatories from around the world.⁸⁶ These signatories are organizations that regularly publish non-partisan reports on the accuracy of statements by public figures, major institutions, and other widely circulated claims of interest to society. It comprises five principles that have to be adhered to and respected by all signatories when conducting their journalistic work. These principles entail the following commitments: non-partisanship and fairness; transparency of sources; transparency of funding and organization; transparency of methodology; open and honest corrections. The aim of this Code is to promote excellence in fact-checking. In order to become a signatory, an extensive accreditation process has to be followed which involves external assessors having to assess the applicant's respect of the Code of Principles based on a checklist.⁸⁷ As explicitly stated in the accreditation process: "priority will be given to applications from countries where Facebook allows third-party fact-checkers to operate".⁸⁸ This is because on 15 December 2016, Facebook declared that adherence to this Code would be the *conditio sine qua non* for becoming a verified fact-checker on its platform when it announced that it would counter 'fake news' by involving third-party verified fact-checkers.⁸⁹ As regards the Netherlands, on 21 September 2017, NUcheckt became a verified signatory, followed by Nieuwscheckers on 22 November 2017.⁹⁰
- The EBU: Representing 73 public service media organizations across 56 countries, the EBU provides an overview of the fact-checking initiatives developed or supported by its

⁸⁵ See, in this connection, the Appendix to the 'Information Disorder' study, entitled, 'European Fact-checking and Debunking Initiatives', pp. 86 *et seq.*

⁸⁶ Poynter, 'International Fact-Checking Network fact-checkers' code of principles', available at: <https://www.poynter.org/international-fact-checking-network-fact-checkers-code-principles>

⁸⁷ Poynter, 'Application process for signatories of the Fact-Checkers' Code of Principles', available at: <https://docs.google.com/document/d/1TVH6Xduaz8IYxvnRfMzi85PMTxCseNoUQ-gcdqlsnoI/edit>

⁸⁸ *Ibid.*, see Appendix.

⁸⁹ Benjamin Mullin and Alexios Mantzarlis, 'Facebook has a plan to fight fake news. Here's where we come in', Poynter, 15 December 2016, available at: <https://www.poynter.org/news/facebook-has-plan-fight-fake-news-heres-where-we-come>.

⁹⁰ Poynter, 'International Fact-Checking Network fact-checkers' code of principles', *op. cit.*

members in the Annex to its *Position Paper: “Fake news” and the information disorder*.⁹¹

- Google’s Fact-check label: Google has adopted fact-checking measures by providing a ‘fact-check label’ on both Google News and Google Search, through which news publishers or fact-checking organisations can indicate which articles have been fact-checked by them. This label is then displayed on the search results page, in the snippet of the identified article. In April 2017, Google made this feature available globally.⁹² Google News Lab, an organisation separate from Google News, also collaborates with journalists and other media actors on fact-checking amongst other activities.⁹³
- The InVID innovation: This action is funded by the European Commission. It consists in the creation of a platform that will provide services to detect, authenticate and check the reliability and accuracy of newsworthy video files and video content spread via social media. InVID released a verification plug-in in July 2017, which was conceived as a toolbox to help journalists verify images and videos as well as debunk disinformation.⁹⁴ The toolbox has since been adopted by hundreds of active users from the main media newsrooms, NGOs dealing with human rights and media educators.⁹⁵
- ‘CrossCheck’: In February 2017, Google and First Draft launched ‘CrossCheck’, an online collaborative journalism project between news, education and technology partners. Its first project was aimed at helping French voters obtain reliable and truthful information during the ten weeks leading up to the French presidential election, in order to vote in an informed way. This brought together 37 newsroom partners in France and the UK to help report false information. Social media posts, online news articles, videos, promotional images and campaign statements were investigated by participating news partners in order to check their accuracy. The public at large was also involved in this project. It could, by means of an online form, notify the team about suspicious articles. The evidence relied on by the journalist in order to reach a conclusion, as well as every

⁹¹ EBU, Position Paper: “Fake news” and the information disorder, 18 April 2018.

⁹² HLEG Final Report, p. 14 (footnote 9) - See Google Blog post: <https://www.blog.google/products/search/fact-check-now-available-google-search-and-news-around-world/>

⁹³ Mentioned in ‘Information Disorder’ study, p. 59, Google News Lab website: <https://newsinitiative.withgoogle.com/training/>

⁹⁴ See: <http://www.invid-project.eu/tools-and-services/invid-verification-plugin/>

⁹⁵ HLEG Final Report, p. 16 (footnote 14).

step of his/her investigation, had to be carefully reported and shared with journalists from other partner organisations. A conclusion would only be published once the investigation had been checked and validated by at least one other participating news partner.⁹⁶

- Faktisk: Faktisk is a Norwegian ‘fact-checking’ website, launched in July 2017 by Norway’s four biggest news organisations. It is a non-profit organisation but receives more than 50 percent of its annual budget from its partners. It is thus privately financed. The fact-check is displayed by means of a five-point scale ranging from absolutely true to absolutely false. Once fact-checked, the fact checks can be embedded by everyone in their own work. Faktisk has proven to be a big success and has been ranked among Norway’s most popular sites.⁹⁷
- Veriflix: Veriflix is an initiative of Roularta Media group, in collaboration with KU Leuven University and the start-up, Look Live Media. It aims to develop a set of tools in order for news media organisations to counter ‘fake news’ stemming from user-generated videos.⁹⁸ This will be done with the help of artificial intelligence. Accordingly, “based on several key components, publishers will be able to drive long-term value for their audiences due to the distribution of authentic, fact-checked, quality content related to online news stories”.⁹⁹ In December 2017 and as part of its ‘Digital News Initiative Innovation Fund’, Google decided to support this project with €400,000.¹⁰⁰
- RTL Group’s “fake news verification team”: RTL Group, the leading European entertainment network, has set up an internal “fake news verification team”, consisting of international experts from across RTL Group. This team is responsible for verifying the authenticity of news content. Research results are immediately shared throughout the Group which allows it to counter false information at an international level. The team

⁹⁶ ‘Information Disorder’ study, p. 66; HLEG Final Report, p. 15 - Crosscheck site:

<https://crosscheck.firstdraftnews.org/france-en/>

⁹⁷ See Faktisk site: <https://www.faktisk.no/> and information about the site can be found here:

<https://www.faktisk.no/>

⁹⁸ Innovation Fund, Veriflix,

⁹⁹ *Ibid.*

¹⁰⁰ Knack.be, ‘Google geeft Roularta 400.000 euro om fake news te detecteren’, 18 December 2017, available at: <https://datanews.knack.be/ict/google-geeft-roularta-400-000-euro-om-fake-news-te-detecteren/article-normal-940857.html>.

regularly offers a variety of trainings, aimed at verifying the authenticity of information, to all editorial teams within the Group.¹⁰¹

4.2.2 Flagging / labelling / blacklisting

Flagging, labelling and blacklisting are all means through which content or content creators can be marked as constituting disinformation, or as otherwise harmful/untruthful. Some online platforms allow users to flag posts as fake or false, but there are also independent organisations working on flagging or blacklisting solutions. The main aim of these measures is to identify and signal content that is (potentially) false and thereby, more generally, raise awareness of such content. In some cases, for example, in the case of blacklisting, the purpose may go beyond mere awareness-raising and take on a warning function. Some flagging and labelling initiatives may include verification and correction features (akin to fact-checking), but they do not necessarily engage in those activities.

Examples:

- ‘Check This’ is a Google Chrome browser extension that gives additional information on news sources.¹⁰² The extension indicates, for instance, when news is fake, satire, far-left or far-right. The extension presents this information in the form of a pop-up box. The goal is to raise awareness, which in turn allows citizens to make better informed choices about the news they consume.
- YouTube’s Trusted Flagger program was set up “to help provide robust tools for government agencies and non-governmental organizations (NGOs) that are particularly effective at notifying YouTube of content that violates” its Community Guidelines.¹⁰³ Videos flagged under this program are subsequently reviewed by YouTube’s content moderators. The program is limited to the flagging of content that infringes YouTube’s Community Guidelines. It is not intended for the flagging of illegal content for removal, which is a separate procedure that involves the submission of a content removal form.

¹⁰¹ HLEG Final Report, p. 15 (footnote 11) - See Backstage Report RTL Group: https://backstage.rtlgroup.com/en/news/2017/6/alliance_against_fake_news.cfm.

¹⁰² ‘Information Disorder’ study, p. 63 - Google Store link to browser extension: <https://chrome.google.com/webstore/detail/check-this-by-metacert/felmjclcjadopolhjmlbemfekjaojfbn?hl=en>

¹⁰³ For details, see: <https://support.google.com/youtube/answer/7554338?hl=en>.

- The French newspaper *Le Monde* has a tool on its website which allows users to search for a URL and see whether the website has been associated with unreliable content in the past.¹⁰⁴ The tool will advise caution in case of questionable websites, and back this advice up with references. If a website has disseminated false or misleading articles in the past, the tool advises people to try to ascertain the source of the information.
- In 2015, the EU's European External Action Service East Stratcom Task Force started the EU vs Disinformation campaign and website.¹⁰⁵ Its aim is to forecast, address and respond to primarily pro-Kremlin disinformation. In the Netherlands, the campaign was criticized because three Dutch media outlets were (incorrectly) labelled as having disseminated disinformation. This led the outlets to demand an explanation and launch civil legal action against the EU.¹⁰⁶ The case was subsequently withdrawn after the articles in question were removed from the EU vs Disinformation website and an apology was made.¹⁰⁷ In March 2018, the Lower House of the Dutch Parliament supported a motion to abolish EU vs Disinformation.¹⁰⁸
- The Open Brand Safety framework is an effort to create a comprehensive list of 'fake news' websites, for the purpose of dissuading advertisers from promoting their products on those websites.¹⁰⁹ This stops money from flowing to unreliable news websites. The Open Brand Safety framework is a collaborative effort by social news agency Storyful, advertising analytics firm Moat and CUNY Journalism School.

4.3 Containing or corrective measures

4.3.1 Contextualization of (fake) news

Contextualisation measures provide additional information and context in order to demonstrate the falsity, inaccuracy or incompleteness of particular news content. Research-

¹⁰⁴ 'Information Disorder' study, p. 63 - *Le Monde* website: <http://www.lemonde.fr/verification/>.

¹⁰⁵ See: <https://euvsdisinfo.eu/>.

¹⁰⁶ NOS, 'GeenStijl, TPO en Persgroep slepen EU voor de rechter', 20 February 2018, available at: <https://nos.nl/artikel/2218452-geenstijl-tpo-en-persgroep-slepen-eu-voor-de-rechter.html>.

¹⁰⁷ NOS, 'Media trekken kort geding tegen EU om nepnieuws in', 13 March 2018, available at: <https://nos.nl/artikel/2222185-media-trekken-kort-geding-tegen-eu-om-nepnieuws-in.html>.

¹⁰⁸ See: <https://nos.nl/artikel/2220926-tweede-kamer-stemt-voor-opheffing-nepnieuwsbureau-eu.html> and <https://nos.nl/artikel/2221353-ook-ollongren-wil-nu-af-van-europese-waakhond-nepnieuws.html>.

¹⁰⁹ 'Information Disorder' study, p. 63; Nieman Lab article: <http://www.niemanlab.org/2017/05/newsonomics-can-a-master-blacklist-choke-off-fake-news-money-supply/>.

driven measures to limit the impact of, or counter the effects of, “fake news” and online disinformation deserve mention. Well-known examples of such (academic) research centres and projects include:

- First Draft: “a project of the Shorenstein Center on Media, Politics and Public Policy at Harvard University’s John F. Kennedy School of Government – uses research-based methods to fight mis- and disinformation online. Additionally, it provides practical and ethical guidance in how to find, verify and publish content sourced from the social web”.¹¹⁰
- The LSE Truth, Trust and Technology (T3) Commission: The T3 Commission “deals with the crisis in public information” and works “with experts, practitioners and the public to identify structural causes of media misinformation and set out a new framework for strategic policy”.¹¹¹
- Data & Society Disinformation Action Lab (DAL): The DAL works with a range of partners to attempt to find “innovative approaches to addressing some the complex dynamics underpinning the spread of propaganda and disinformation”.¹¹² Data & Society “is a research institute in New York City that is focused on the social and cultural issues arising from data-centric and automated technologies”.¹¹³

Another one of the main contextualisation strategies is to point readers/viewers of news towards alternative or opposing sources, with the aim of exposing them to a wider range of sources and viewpoints. On 20 December 2017, Facebook announced that it would make use of a ‘Related Articles’ tool, enabling it to contextualise content and thereby counter fake news. This method involves Facebook automatically recommending a few alternative news sources covering the same topic below news articles on heavily discussed topics. The decision to do so and to leave behind the disputed flagging method is, as was stated in Facebook’s announcement, based on academic research. According to some research, flagging may be counter-productive and lead to undesired results (see further, below), whereas the contextualization of posts enables users to better go to the facts and is more effective. Indeed, Facebook has found out that when displaying related articles fewer shares occur by users than when the Disputed Flag is shown. Moreover, this permits quicker results

¹¹⁰ <https://firstdraftnews.org/about/>.

¹¹¹ <http://www.lse.ac.uk/media-and-communications/truth-trust-and-technology-commission>.

¹¹² <https://datasociety.net/research/disinformation-action-lab/>.

¹¹³ <https://datasociety.net/about/>.

as only one fact checker is necessary for this new tool, whereas for flagging at least two fact-checkers were required.¹¹⁴

4.3.2 Using ‘nudges’ to slow sharing

Reddit’s largest news community, /r/worldnews, has had success by using ‘nudges’ to increase fact-checking and decrease sharing.¹¹⁵ Their research showed that posting a disclaimer about fact-checking and alternative sources under articles written by websites that had received a significant amount of complaints decreased the total amount of upvotes such articles got. In posts that had the ‘stickied’ disclaimer, users were encouraged to do some research into the topic of the article before posting comments and therefore took part in significantly less divisive discourse. This method therefore provided a way to mitigate the effects of false or misleading news without applying censorship or restricting users’ behaviour.

4.3.3 Closing automated accounts

Two of the biggest tech companies have targeted large networks of automated accounts in the past in an effort to reduce the spread of fake or misleading information on their platforms. Facebook deleted tens of thousands of accounts before the French and UK elections.¹¹⁶ In order to do so and to thereby identify unauthentic accounts, the company has incorporated a system which looks for patterns, such as the repeated posting of the same content or a sudden increase in the account’s messaging activity.¹¹⁷ One of the criticisms of this practice is that Facebook should not limit such measures to pre-election times. In June 2017, Twitter updated

¹¹⁴ Tessa Lyons, ‘Replacing Disputed Flags with Related Articles’, 20 December 2017, <https://newsroom.fb.com/news/2017/12/news-feed-fyi-updates-in-our-fight-against-misinformation/>; HLEG Final Report, p. 15, footnote 14; ‘Information Disorder’ study, p. 60; Academic article on this: Leticia Bode and Emily K. Vraga, ‘In Related News, That Was Wrong: The Correction of Mis-information Through Related Stories Functionality in Social Media’, 65 *Journal of Communication*, (Issue 4, August 2015): 619-638. (‘Information Disorder’ study, fn. 175).

¹¹⁵ ‘Information Disorder’ study, p. 60 - Nieman Lab Article: <http://www.niemanlab.org/2017/02/reddits-worldnews-community-used-a-series-of-nudges-to-push-users-to-fact-check-suspicious-news/>.

¹¹⁶ ‘Information Disorder’ study, p. 61 - Rappler article: <https://www.rappler.com/newsbreak/investigative/148347-fake-accounts-manufactured-reality-social-media>.

¹¹⁷ TechCrunch, ‘Facebook culls tens of thousands of fake accounts ahead of UK election’, 8 May 2017, <https://techcrunch.com/2017/05/08/facebook-culls-tens-of-thousands-of-fake-accounts-ahead-of-uk-election/>; NBCnews, ‘Facebook shuts down 30 000 fake Accounts in France ahead of Presidential Elections’, 15 April 2017, <https://www.nbcnews.com/news/world/facebook-shuts-down-30-000-fake-accounts-france-ahead-presidential-n746821>.

the section ‘Automation Rules’ under its rules and policies in order to better regulate and counter automated accounts and tweets.¹¹⁸

4.4 Regulatory and non-regulatory counter measures

4.4.1 Regulatory measures

As explained above, caution is called for regarding any proposals to regulate “fake news”, lest such regulation would breach European legal standards guaranteeing the right to freedom of expression. Recent legislative initiatives in European countries have drawn considerable criticism on account of their (likely) implications for freedom of expression.

4.4.1.1 Germany

The controversial German Act to Improve Enforcement of the Law in Social Networks¹¹⁹ creates new, heavy obligations for social media platforms to promptly deal with, and to report complaints about, unlawful content.¹²⁰ Such content could include particular types of “fake news”, but only insofar as they are covered by existing provisions in German criminal law.¹²¹ Although the Act does not explicitly mention “fake news”, its explanatory memorandum does refer to “fake news” as being one of the reasons for such legislation.¹²²

The general rule, covered by Section 3 of the Act, is that manifestly unlawful content shall be blocked or removed within 24 hours after having been notified. In other cases, where the

¹¹⁸ ‘Information Disorder’ study, p. 62, Twitter rules update: <https://help.twitter.com/en/rules-and-policies/twitter-automation>.

¹¹⁹ Act to Improve Enforcement of the Law in Social Networks (Network Enforcement Act)/*Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken (Netzwerkdurchsetzungsgesetz – NetzDG)*, 1 September 2017, *Bundesgesetzblatt*, Jahrgang 2017, Teil I, Nr. 61, 7 September 2017 (entry into force: 1 October 2017).

¹²⁰ For academic analysis, see: Gerald Spindler, ‘Internet Intermediary Liability Reloaded – The New German Act on Responsibility of Social Networks and its (In-) Compatibility with European Law’, 8 (2017) *JIPITEC* 166.

¹²¹ See further: Staff and agencies, ‘Germany approves plans to fine social media firms up to €50m’, *The Guardian*, 30 June 2017, available at: <https://www.theguardian.com/media/2017/jun/30/germany-approves-plans-to-fine-social-media-firms-up-to-50m>.

¹²² See Explanatory Memorandum (Section II) of Act to Improve Enforcement of the Law in Social Networks (Network Enforcement Act)/*Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken (Netzwerkdurchsetzungsgesetz – NetzDG)*, 1 September 2017, *Bundesgesetzblatt*, Jahrgang 2017, Teil I, Nr. 61, 7 September 2017 (entry into force: 1 October 2017).

illegality is less apparent, a 7-day time limit applies.¹²³ When “the decision regarding the unlawfulness of the content is dependent on the falsity of a factual allegation or is clearly dependent on other factual circumstances”, the 7-day time limit may be exceeded.¹²⁴

When a platform receives more than 100 complaints about unlawful content per calendar year, a transparency obligation applies consisting of drawing up reports and publishing these in the Federal Gazette and on the platform’s website.¹²⁵ Where a platform fails to comply with its obligations, fines ranging from five hundred thousand Euros to five million Euros can be imposed.¹²⁶ Importantly, the different obligations only apply to social network platforms having more than two million registered users in the Federal Republic of Germany.¹²⁷ Platforms that are explicitly excluded from the law’s scope of application are those which offer journalistic or editorial content and are responsible for the displayed/produced content, as well as those “which are designed to enable individual communication or the dissemination of specific content”.¹²⁸

The Act has been criticized for its chilling effect on free speech.¹²⁹ Its critics argue that it puts responsibility for the determination of complicated questions of legality in the hands of private actors.¹³⁰ Entrusting that task to private actors is problematic from a rule of law perspective. Moreover, the use of overbroad terms, coupled with the risk of serious fines, is likely to lead to over-removal of content. As Bernd Holznagel stated in his legal analysis on the draft version of the Law, written on behalf of the OSCE Representative on Freedom of the Media: “with the risk of high fines in mind, the networks will probably be more inclined to delete a post than to expose themselves to the risk of a penalty payment. As the differentiation between “unlawful” and “manifestly unlawful” is anything but clear, networks will, in case of doubt, probably erase contributions”.¹³¹ Moreover, no concrete guidance is given to platforms about how to balance the authors’ interests against the interests of the

¹²³ Article 1, Section 3(3).

¹²⁴ Article 1, Section 3(3) sub (a).

¹²⁵ Article 1, Section 2.

¹²⁶ Article 1, Section 4.

¹²⁷ Article 1, Section 1.

¹²⁸ *Ibid.*

¹²⁹ See, for example, ARTICLE 19, ‘Germany: The Act to Improve Enforcement of the Law in Social Networks’, Legal Analysis, August 2017, available at: <https://www.article19.org/data/files/medialibrary/38860/170901-Legal-Analysis-German-NetzDG-Act.pdf>.

¹³⁰ Human Rights Watch, ‘Germany: Flawed Social Media Law NetzDG is Wrong Response to Online Abuse’, 14 February 2018, available at: <https://www.hrw.org/news/2018/02/14/germany-flawed-social-media-law>.

¹³¹ Dr. Bernd Holznagel, ‘Legal Review of the Draft Law on Better Law Enforcement in Social Networks’, OSCE RFOM, May 2017, p. 23, available at: <https://www.osce.org/fom/333541?download=true>.

notifying users.¹³² Such fears are very much in line with those expressed by David Kaye, UN Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, who claimed that the bill was, for many reasons, not in accordance with international human rights law.¹³³ European Digital Rights (EDRi) has criticized the absence of a provision for the right of users to appeal against the deletion of content.¹³⁴ Human Rights Watch has also been critical of this lack of judicial oversight or judicial remedies.¹³⁵

4.4.1.2 Italy

In Italy, in February 2017, a Senator introduced a legislative proposal which aimed to create specific new provisions to criminalize different types of conduct relating to the dissemination of “fake news”.¹³⁶ Under the proposed new provision, anyone who “publishes or spreads via the Internet, fake news or exaggerated or biased information on manifestly ill-founded or false facts and circumstances shall be punished by a fine of up to EUR 5,000”.¹³⁷ The proposal was not adopted.¹³⁸

However, on 18 January 2018, the Minister of the Interior introduced the “Operating Protocol for the Fight Against the Diffusion of Fake News through the Web on the Occasion of the Election Campaign for the 2018 Political Elections” on 18 January 2018. The Operating Protocol provided for a “red button” reporting service whereby users “may indicate the existence of a network of content attributable to fake news” (“the portal”). *Polizia Postale*, a unit of the Italian State Police that investigates cyber-crime, was tasked with reviewing reports and act accordingly.

The United Nations Special Rapporteur on Freedom of Opinion and Expression, David Kaye, has expressed his concern that the Operating Protocol and its implementation, “coupled with

¹³² *Ibid.*

¹³³ David Kaye, ‘Mandate of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression’, Reference: OL DEU 1/2017, 1 June 2017, available at: <http://www.ohchr.org/Documents/Issues/Opinion/Legislation/OL-DEU-1-2017.pdf>.

¹³⁴ EDRi, Recommendations on the German bill “Improving Law Enforcement on Social Networks” (NetzDG), 20 June 2017, available at: https://edri.org/files/consultations/tris_netzdg_edricontribution_20170620.pdf.

¹³⁵ Human Rights Watch, ‘Germany: Flawed Social Media Law’, *op. cit.*

¹³⁶ Ernesto Apa and Marco Bassini, ‘Italy: Legislative proposal on fake news’, *IRIS – Legal Observations of the European Audiovisual Observatory*, 2017-5/27, available at: <https://merlin.obs.coe.int/iris/2017/5/article27.en.html>.

¹³⁷ *Ibid.*

¹³⁸ UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Letter to the Italian authorities, 20 March 2018, p. 1.

the threat of criminal sanctions”, are problematic from the perspective of freedom of expression.¹³⁹ He recalled the relevance of Article 19 ICCPR and the Joint Declaration on freedom of expression and “fake news” of 2017. The latter concluded that “general prohibitions on the dissemination of information based on vague and ambiguous ideas, including “false news” or “non-objective information” are incompatible and should be abolished”.¹⁴⁰ The Protocol aims to combat “manifestly unfounded and biased news, or openly defamatory content”; the Special Rapporteur criticised the undefined and vague nature of these terms.¹⁴¹ The background of Italy’s existing (criminal) defamation laws, to which the Protocol is linked, was also identified as another cause for concern.¹⁴² The Special Rapporteur urged the Italian Government to “consider alternative measures such as the promotion of independent fact-checking mechanisms, public education and media literacy, which have been recognized as less intrusive means to address disinformation and propaganda”.¹⁴³

Another recent development in Italy is noteworthy: the publication by the Italian Communication Authority (AGCOM) of self-regulatory guidelines to ensure equal treatment of parties/candidates on online platforms within the context of the (then) pending general elections.¹⁴⁴ The guidelines deal with issues such as: equal treatment of political subjects; transparency of online political advertising; reporting mechanisms for illegal content (opinion polls in the 15 days preceding election day and defamation of electoral candidates); public entities refraining from using social media for political communication during election periods; the “desirability” of platforms preventing political groups from campaigning on the eve and day of elections, and the enhancement by Facebook and Google of existing fact-checking mechanisms.¹⁴⁵

4.4.1.3 France

¹³⁹ *Ibid.*, p. 4.

¹⁴⁰ *Ibid.*, p. 3.

¹⁴¹ *Ibid.*, p. 4.

¹⁴² *Ibid.*, p. 4.

¹⁴³ *Ibid.*, p. 5.

¹⁴⁴ AGCOM, Guidelines for equal access to online platforms during the election campaign for the 2018 general elections, 1 February 2018.

¹⁴⁵ For more details and analysis, see: Ernesto Apa and Filippo Frigerio, ‘Italy: Self-regulatory guidelines for online platforms for next general elections published by Italian Communication Authority’, *IRIS* 2018-3: 24, available at: <https://merlin.obs.coe.int/iris/2018/3/article24.en.html>.

In France, President Macron has announced his intention to regulate social media in respect of, amongst other things, the dissemination of “fake news” during election periods.¹⁴⁶ The draft law proposal was leaked on 7 March by the site Nextinpact,¹⁴⁷ shortly after having been consulted and discussed by newspaper *Lemonde.fr*.¹⁴⁸ However, on 16 March 2018 and on 21 March 2018, the law was officially filed at the French National Assembly.¹⁴⁹ On the former date, the organic law was filed and on the latter date it was the ordinary law’s turn. The need for those two laws is regulated under Article 6 of the French Constitution which stipulates that the rules concerning the President’s election are to be determined by organic law.¹⁵⁰ Therefore, since the law against fake news touches upon election times, such an organic law was necessary in order for the ordinary law against fake news to take effect.¹⁵¹

Under Title I of the law, platforms are obliged to be transparent about sponsored content, by revealing both the identity of the advertiser and the persons who control it or on whose behalf it acts, as well as the funds devoted to the promotion of such content (only when those funds are higher than a certain threshold to be defined by decree).¹⁵² This obligation would only apply during election periods (which means from the date of publication of the order calling the election until the end of voting operations) and only to news related information and therefore not to content aimed at promoting goods or services.¹⁵³ Moreover, only platforms whose activity exceeds a threshold number of connections on the French territory will be subject to this obligation.¹⁵⁴ Failing to obey to these obligations might lead to sanctions of one year’s imprisonment and a €75,000 fine.

¹⁴⁶ See further: <https://www.theguardian.com/world/2018/jan/03/emmanuel-macron-ban-fake-news-french-president>.

¹⁴⁷ Nextinpact, ‘Exclusif, téléchargez la future loi contre les Fake News’, 7 March 2018, available at: <https://www.nextinpact.com/news/106262-exclusif-telecharger-future-loi-contre-fakes-news.htm>.

¹⁴⁸ LeMonde.fr, ‘«Fake News » : Les pistes du texte de loi en préparation’, 7 March 2018, available at: http://www.lemonde.fr/actualite-medias/article/2018/03/07/fake-news-les-pistes-du-texte-de-loi-en-preparation_5266947_3236.html.

¹⁴⁹ Assemblée Nationale, n. 799 - ‘Proposition de Loi relative à la lutte contre les Fausses Informations’, 21 March 2018, available at: <http://www.assemblee-nationale.fr/15/propositions/pion0799.asp>; Assemblée Nationale, n. 772 - ‘Proposition de Loi organique relative à la lutte contre les Fausses informations’, 16 March 2018, available at: <http://www.assemblee-nationale.fr/15/propositions/pion0772.asp>.

¹⁵⁰ Article 6 of the French Constitution, available at: <https://www.legifrance.gouv.fr/affichTexteArticle.do?idArticle=LEGIARTI000019241002&cidTexte=LEGITEXT000006071194&dateTexte=20140401>.

¹⁵¹ Loi organique relative à la lutte contre les fausses informations, available at: <https://cdn2.nextinpact.com/medias/fake-news-organique---03-2017.pdf>.

¹⁵² Article 1er (I) sub 2, Assemblée Nationale, n. 799 - ‘Proposition de Loi relative à la lutte contre les Fausses Informations’, 21 March 2018, available at: <http://www.assemblee-nationale.fr/15/propositions/pion0799.asp>.

¹⁵³ *Ibid.*, Explanatory memorandum, p. 2.

¹⁵⁴ *Ibid.*, Article 1er (I) sub 2.

Title I also introduces a judicial interim procedure aimed at quickly stopping the dissemination of false news during election periods. Consequently, an interim proceeding could be started (the Paris Regional Court will have exclusive jurisdiction) at the request of the public prosecutor or any other party responsible for taking action as soon as false information likely to alter the integrity of the poll has been disseminated online, in a massive and artificial manner and in particular via the use of automated “bots”. Once seized, the interim relief judge will have 48 hours to decide about the case and may order the site diffusing “fake news” to be delisted or the fake content to be removed together with the banning of its return online. He may also oblige the closure of a user’s account who has repeatedly contributed to the dissemination of the fake content or order the blocking of access to the website.¹⁵⁵ Hosters, platforms and Internet access providers will be required to comply with the decision under penalty of sanctions. Importantly, the law does not define “fake news” as it is already defined in Article 27 of the freedom of the press law,¹⁵⁶ but the explanatory memorandum explicitly mentions that parody or satirical content are excluded from the law’s scope of application.¹⁵⁷

Under Title II, the powers of the *Conseil supérieur de l’audiovisuel* (CSA, the national audiovisual regulator), are strengthened as it will have the possibility “to prevent, suspend or terminate” the broadcasting of television services controlled by a foreign State, undermining the higher interests of the French Nation or participating in an undertaking to destabilize the nation’s institutions. Furthermore, under Title III, the law would impose a new “duty of cooperation” on social networks and Internet service providers, which consists of three measures and which would not be restricted to electoral periods. First, they would have to offer any Internet user an easily accessible and visible way to report “fake news”. Second, they must “promptly inform the competent public authorities of any activity to disseminate such false information brought to their attention”. Finally, they must make public the means they devote to the fight against the dissemination of false information.¹⁵⁸

¹⁵⁵ *Ibid.*, Article 1er (I) sub 2 jo. Explanatory Memorandum, p. 2.

¹⁵⁶ Loi du 29 Juillet 1881 sur la liberté de la presse, Article 27, available at: <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006070722>.

¹⁵⁷ Explanatory memorandum, p. 2, Assemblée Nationale, n. 799 - ‘Proposition de Loi relative à la lutte contre les Fausses Informations’, 21 March 2018, available at: <http://www.assemblee-nationale.fr/15/propositions/pion0799.asp>.

¹⁵⁸ *Ibid.*, Article 9.

Importantly, the Council of State will have to deliver an advice about this law, which means that the text might still be subject to change. Moreover, since such a law affects the information society, it will first have to be notified to the European Commission,¹⁵⁹ similarly to the German Network Enforcement Act.¹⁶⁰ If it is adopted in the coming period, this law could be applied for the first time during the 2019 European election campaign.¹⁶¹

4.4.2 Non-regulatory counter measures

4.4.2.1 Credibility scores

Credibility scoring is the exercise of awarding news content a certain score for the purpose of up- or down-ranking it. Poor quality content that is, for instance, false or misleading would therefore show up in people's news feed less often. The end goal for projects trying to define a set list of criteria for rating news content for credibility is to be integrated into Facebook/Google algorithms and help shape the way news is presented online.

Examples:

- Trust scores based on users' surveys: Facebook has announced the development of trust scores based on users' quality surveys. Users will be asked to indicate whether they know the source of a certain post, and if so, whether they trust it. According to Facebook, the users' community is the most objective method for determining which news sources are most broadly trusted. Based on the surveys and a change in algorithms, users' news feeds will be filled with news originating from the media that are most familiar to them and experienced by the public as being "credible".¹⁶²

¹⁵⁹ Directive (EU) 2015/1535 of the European Parliament and of the Council of 9 September 2015 laying down a procedure for the provision of information in the field of technical regulations and of rules on Information Society services (Text with EEA relevance), see 'The notification procedure in brief', available at: <http://ec.europa.eu/growth/tools-databases/tris/en/about-the-20151535/the-notification-procedure-in-brief/>.

¹⁶⁰ European Commission, 'Notification Detail – Act improving law enforcement on Social Networks', 28 June 2017, available at: <http://ec.europa.eu/growth/tools-databases/tris/en/search/?trisaaction=search.detail&year=2017&num=127>.

¹⁶¹ RT France, 'Le gouvernement a transmis sa loi contre les "Fake News" aux députés LREM... mais elle a déjà fuité', 7 March 2018, available at: <https://francais.rt.com/france/48558-projet-loi-gouvernement-contre-fake-news-communique-lrem-deja-fuite>.

¹⁶² HLEG Final Report, p. 16 (footnote 13) - See: <https://www.nytimes.com/2018/01/19/technology/facebook-news-feed.html>.

- ‘NewsCheck’: In January 2017, First Draft launched ‘NewsCheck’, a Google Chrome browser extension which allows its users to see how a video or image has been authenticated, in terms of trustworthiness, by others. Through this tool, the Visual Verification Guides, earlier developed by First Draft, are implemented in a user-friendly way. Users wanting to make use of this tool will have to carry out a four-step checklist test, consisting of the following questions: ‘Are you looking at the original version? Do you know who captured the photo? Do you know where the photo was captured? Do you know when the photo was captured?’ Each question can be answered by means of a colour-coded, pre-defined answer ranging from green (trustworthy) to red (not trustworthy). After having answered those four questions by means of colour-ranked answers, a non-scientific score is shown to the user which represents the level of trustworthiness of the post. This in turn permits users to embed the video/image on their own website. In doing so, the verification results of the checklist are placed alongside the embedded image/video in order for other users to understand how trustworthy it is.¹⁶³ Such a tool is interesting as it shows that some indicators, worked out in checklists, are being developed in order to calculate how trustworthy a piece of content is. Based on this, different projects have seen the light of day, such as Misinfocon’s plan to build technical standards for credibility¹⁶⁴ or the Trust project at Santa Clara University aimed at developing Trust indicators.¹⁶⁵

Google Search’s Quality raters: Indirectly related to credibility scores, is Google’s use of ‘Search quality raters’ - people with whom Google contracts, who are given the task to evaluate Google’s search results in order to find out whether the results meet users’ expectations.¹⁶⁶ They do so by rating the quality of pages that appear in Google’s top results, based on a set of Guidelines,¹⁶⁷ which Google in turn uses to improve its search

¹⁶³ First Draft, ‘Launching our new Chrome extension NewsCheck’, 12 January 2017, available at: <https://firstdraftnews.org/launching-new-chrome-extension-newscheck/>.

¹⁶⁴ ‘Information Disorder’ study, p. 64 - website: <https://misinfocon.com/building-technical-standards-for-credibility-59ef9ee4ab73>; An Xiao Mina, ‘Building Technical Standards for Credibility’, 3 March 2017, available at: <https://misinfocon.com/building-technical-standards-for-credibility-59ef9ee4ab73>.

¹⁶⁵ Trust Project website: <https://thetrustproject.org/>; Santa Clara University, ‘Backstory: Trust Indicators Journey’, available at: <https://www.scu.edu/ethics/focus-areas/journalism-ethics/programs/the-trust-project/backstory-trust-indicators-journey/>.

¹⁶⁶ HLEG Final Report, p. 16 (footnote 13).

¹⁶⁷

See: <https://static.googleusercontent.com/media/www.google.com/en//insidesearch/howsearchworks/assets/searchqualityevaluatorguidelines.pdf>.

algorithms. Consequently, the resulting ratings do not directly alter the reviewed website, but help shape the ranking and results systems. In April 2017, Google announced that it had updated its guidelines in order to better tackle fake news and clickbait practices by “providing more detailed examples of low-quality webpages for raters to appropriately flag, which can include misleading information, unexpected offensive results, hoaxes and unsupported conspiracy theories”.¹⁶⁸

4.4.2.2 Changing advertising rules

The underlying business model of many internet platforms encourages the practice of “fake news” by relying on advertisement revenues as main source of income. This opens the door for click-bait practices and misleading advertisements. Consequently, one way to counter fake news is to tackle the heart of the problem, at the level of advertisement.

Examples:

- Preventing revenue from flowing to harmful websites: Facebook and Google have both taken steps to prevent websites that disseminate fake or misleading information from generating revenue through their advertising on their platforms.¹⁶⁹ For example, in 2016, Google suspended more than 1,300 accounts for “tabloid cloaking”, which are misleading ads taking the form of ‘news’ articles that are presented by means of attractive headlines, aimed at generating clicks and leading to articles that have nothing to do with the headlines’ topics.¹⁷⁰ In August 2017, Facebook announced that it had taken measures to prohibit runners of pages that “repeatedly share[s] stories that have been marked as false by third-party fact-checkers” to no longer have the possibility to buy advertisement space on Facebook.¹⁷¹ However, for such a measure to be effective, many more online intermediaries must join in, as reportedly “fake news” creators have recovered from the initial loss of profit by using other advertising networks.¹⁷²

¹⁶⁸ See Google blogpost, ‘Our latest quality improvements for Search’, 25 April 2017, available at: <https://blog.google/products/search/our-latest-quality-improvements-search/>.

¹⁶⁹ ‘Information Disorder’ study, pp. 57-58.

¹⁷⁰ Google blogpost, ‘How we fought bad ads, sites and scammers in 2016’, 25 January 2017, available at: <https://www.blog.google/topics/ads/how-we-fought-bad-ads-sites-and-scammers-2016/>.

¹⁷¹ Facebook News Room, ‘Blocking ads from pages that repeatedly share false news’, 28 August 2017, available at: <https://newsroom.fb.com/news/2017/08/blocking-ads-from-pages-that-repeatedly-share-false-news/>.

¹⁷² ‘Information Disorder’ study, pp. 57-58.

- Warning advertisers of their presence on harmful websites – ‘Sleeping Giants’: Intrinsically related to the issue of advertisement revenue through fake news, is the movement known as the ‘Sleeping Giants’ which started soon after Donald Trump’s presidential election in November 2016 and which aimed at countering fake news by naming and shaming the brands that are advertised on sites divulging fake news, such as *Breitbart News Network*.¹⁷³ In order to do so, these Sleeping Giants make use of Twitter by taking a screenshot of an ad displayed on a site that supports fake news and subsequently bringing it to the attention of the advertiser by means of a tweet accompanied by the tag @slpng_giants.¹⁷⁴ Sites are often not aware that their advertisements are present on fraudulent sites promulgating fake news. Companies can withdraw their advertisements from those undesired websites, which results in those websites making less advertisement revenue and therefore having less financial resources to promulgate “fake news”. Although this movement originated in the US, it has spread across the globe. For example, in March 2017, the Belgian news media *RTLInfo* found out that three Belgian ‘sleeping giants’ were actively taking part in this initiative.¹⁷⁵

¹⁷³ Constellation Research, ‘Sleeping Giants take on Fake News and Brands – Is your Brand Ready?’, 10 January 2017, <https://www.constellationr.com/sleeping-giants-take-fake-news-and-brands-your-brand-ready>

¹⁷⁴ Sleeping Giants, ‘How to be a giant’, 27 November 2016, available at: https://twitter.com/slpng_giants/status/802915214563807232?ref_src=twsrc%5Etfw&ref_url=http%3A%2F%2Fwww.slate.com%2Farticles%2Fnews_and_politics%2Fpolitics%2F2016%2F12%2Fsleeping_giants_campaign_against_breitbart.html&tfw_site=slate.

¹⁷⁵ RTL Info, ‘Des Belges s’organisent pour lutter à leur manière contre les ‘fake news’ qui influencent l’opinion publique : voici leur stratégie’, 13 March 2017, available at: <https://www.rtl.be/info/vous/temoignages/des-belges-s-organisent-pour-lutter-a-leur-maniere-contre-les-fake-news-qui-influencent-l-opinion-publique-voici-leur-strategie-899208.aspx>.

5. Effectiveness of methods/measures against “fake news”

By now, it is very clear that “fake news” and online disinformation can take many forms; are generated and disseminated by different actors, with different motivations, and through different techniques. Chapter 4 has grouped and enumerated a selection of counter-strategies and counter-measures that have been developed by various actors and stake-holders. There is no single panacea for “fake news” and online disinformation. Rather, in light of the complexity of the problem and its multiple dimensions, a range of complementary counter-strategies and counter-measures are called for. Different types of expression that fall under the banner of “fake news” and online disinformation require appropriate, differentiated responses and in some cases no responses at all. The principle of “horses for courses” applies.

The present chapter aims to give a brief and cautious evaluation of the effectiveness of some of the strategies and measures discussed in the previous chapter. One reason for caution is that the evaluation is based on a limited review of relevant literature in a very fast-moving field.¹⁷⁶ Much of the literature surveyed expressly acknowledges that its own findings are limited to specific situations, geographical confines, data-sets, research questions or tests, and that further research is needed before more confident conclusions can be drawn. Calls for further research often emphasize the need for evidence-based and comparative research, as well as for a broadening of research or test focuses from one specific medium or platform to many. There is also a strong case to be made for multi- and inter-disciplinary research owing to the complexity of the field and the relevance of a range of disciplinary perspectives for achieving a rounded understanding of the issues and problems and devising appropriate policy responses.

Another reason for considered caution when evaluating the effectiveness of specific measures and strategies stems from the vagueness of the term “fake news”. As explained at length in Chapter 1, the concept should be studied alongside related concepts, and disaggregated for further analytical precision. The terminological and conceptual framing of research affects its scope and in the absence of widely-accepted and widely-used definitions of relevant terms,

¹⁷⁶ For an extensive overview and analysis of relevant literature, see Tucker *et al.*, ‘Social Media, Political Polarization, and Political Disinformation’, *op. cit.*

different studies may not lend themselves to straightforward comparison. This can make it difficult to meaningfully, or at least accurately, compare research findings.

A third reason for caution is the lack of consensus about the reach and impact, short- and longer-term, of “fake news” and online disinformation. Some recent attempts to measure the reach of these phenomena, for example by the Reuters Institute for the Study of Journalism, have been limited to specific sets of countries (i.c. France and Italy), with the researchers being candidly reluctant to over-extrapolate from their findings.¹⁷⁷ They correctly point out that country-specific media and political contexts are all-important variables in such studies. They also identify commercial and/or political motives as having significant influence over the production and dissemination of disinformation. Another obstacle to wider extrapolation could concern the platforms or means of distribution selected for the research sample; a focus on a particular category for coherence of analysis will not be representative of “the entire wider ecosystem of misinformation and disinformation”.¹⁷⁸ In the case of the Reuters Institute’s study, the focus was on “the 20 most popular false news websites” in France and Italy.¹⁷⁹

The impact of “fake news” and other forms of online disinformation is also a subject of much debate. According to a very recent Flash Eurobarometer on Fake News and Online Disinformation, 83% of the 26,576 European citizens surveyed stated that fake news represent a danger to democracy.¹⁸⁰ Nevertheless, some empirical research is sceptical about the extent of the reach and of the actual impact of “fake news” and disinformation, suggesting that “false news has more limited reach than is sometimes assumed”.¹⁸¹ Similar conclusions have been drawn about filter bubbles, more particularly that “in spite of the serious concerns voiced – at present, there is no empirical evidence that warrants any strong worries about filter bubbles”.¹⁸² Other empirical research has prompted descriptions of “fake news” as

¹⁷⁷ Richard Fletcher, Alessio Cornia, Lucas Graves and Rasmus Kleis Nielsen, ‘Measuring the Reach of ‘Fake News’ and Online Disinformation in Europe’, Reuters Institute for the Study of Journalism, Oxford, 2018.

¹⁷⁸ *Ibid.*, p. 2. See also p. 7, where reference is made to the “potentially ‘long tail’ of false news access”.

¹⁷⁹ *Ibid.*, p. 7.

¹⁸⁰ European Commission, Final results of the Eurobarometer on fake news and online disinformation, Consultation results, 12 March 2018, available at: <https://ec.europa.eu/digital-single-market/en/news/final-results-eurobarometer-fake-news-and-online-disinformation>.

¹⁸¹ Richard Fletcher *et al.*, ‘Measuring the Reach of ‘Fake News’ and Online Disinformation in Europe’, *op. cit.*, p. 7, with additional references to similar independent, evidence-based research findings in the US.

¹⁸² Frederik J. Zuiderveen Borgesius, Damian Trilling, Judith Möller, Balázs Bodó, Claes H. de Vreese and Natali Helberger, ‘Should we worry about filter bubbles?’, 5 *Internet Policy Review* (Issue 1, 2016), p. 10.

“underresearched and overhyped”.¹⁸³ Whereas some people may have dominant or exclusive formal sources of news and information, others are so-called “media omnivores” and have a more varied informational diet.¹⁸⁴ How we shift between use of different media and non-media sources of information can influence how we select, process and be influenced by those sources.¹⁸⁵ It has also been argued that “today, the biggest obstacle to having an informed electorate isn’t fake news but, rather, the ever more commercial, profit-seeking media seeking clicks and eyeballs at the expense of nuance, depth and on-the-ground-reporting”.¹⁸⁶ Others point the finger at the rise and dominance of social media and the lack of funding for public-interest media reporting; factors which have ushered in an era “when everyone has their own facts”.¹⁸⁷

Notwithstanding present gaps in knowledge, understanding and research, the following subsections offer a tentative evaluation of selected strategies and measures for countering “fake news” and online disinformation.

5.1 Preventive or pre-emptive measures

5.1.1 Media, information and news literacy

As media literacy is concerned with enabling individuals of all ages and from all walks of life to acquire and develop the varied set of skills that they need to effectively participate in the present media ecosystem, it will always be a “work-in-progress”. Media literacy can therefore be seen as a policy goal of conduct rather than result; what is important is that effective measures towards a goal are taken, thereby ensuring the progressive realisation of the goal.

¹⁸³ William H. Dutton, ‘Fake news, echo chambers and filter bubbles: Underresearched and overhyped’, *The Conversation*, 5 May 2017, based on William H. Dutton, Bianca C. Reisdorf, Elizabeth Dubois and Grant Blank, ‘Social shaping of the politics of Internet search and networking: Moving beyond filter bubbles, echo chambers, and fake news’, Quello Center Working Paper No. 2944191, SSRN, 31 March 2017.

¹⁸⁴ William H. Dutton, ‘Fake news, echo chambers and filter bubbles: Underresearched and overhyped’, *op. cit.*

¹⁸⁵ Jesper Strömbäck, Kajsa Falasca and Sanne Kruikemeier (2018), ‘The Mix of Media Use Matters: Investigating the Effects of Individual News Repertoires on Offline and Online Political Participation’, *Political Communication* 35:3, 413-432.

¹⁸⁶ Cas Mudde, ‘Why the hysteria around the ‘fake news epidemic’ is a distraction’, *The Guardian*, 7 February 2018, available at: <https://www.theguardian.com/commentisfree/2018/feb/07/hysteria-fake-news-epidemic-distraction>.

¹⁸⁷ Katharine Viner, ‘How technology disrupted the truth’, *The Guardian*, 12 July 2016, available at: <https://www.theguardian.com/media/2016/jul/12/how-technology-disrupted-the-truth>.

In the field of media literacy, there is considerable terminological divergence: media literacy, (critical) information literacy, digital literacy, news literacy, search engine literacy, etc., are all conceptually congruent.¹⁸⁸ It is important not to be distracted by the divergence in terminology and to focus instead on the congruence of relevant goals and processes. Various authors suggest that media literacy suffices as a generic term as it offers a more “holistic” perspective.¹⁸⁹ National media literacy policies, which connect relevant players and stakeholders and facilitate and promote collaborative and individual initiatives, can be very important for the operationalisation of media literacy. Likewise, it is very important to recognise that a multi-actor approach is called for, with different actors (i.e., State bodies, the media, educators, civil society, individuals, Internet service providers, etc.) playing different roles.

5.1.2 Trust-enhancing practices

Individual and public trust in the media and new media actors is not a given – it must be earned and maintained. Initiatives by media actors and organisations to raise their game in terms of professional standards and ethics can help to (re-)gain public trust and confidence.

Disillusion and declining trust can widen the distance between audiences/readers and the media, old and new. That distance can become fertile ground for disinformation to grow. While initiatives like the EJM (and probably the JTI too, in due course) have developed strong profiles and are developing capacity and resources that are increasingly widely used, other actors have roles and responsibilities too, and markets can also have unpredictable effects on the sustainability of accountable and quality journalism.

In its Resolution, ‘The protection of editorial integrity’, the Parliamentary Assembly of the Council of Europe (PACE) has expressed its support for the work of various organisations to enhance ethical practices in the media, namely: the European Federation of Journalists, the EBU, the Alliance of Independent Press Councils of Europe and the EJM.¹⁹⁰

¹⁸⁸ Tarlach McGonagle, ‘Media Literacy: No Longer the Shrinking Violet of European Audiovisual Media Regulation?’, in S. Nikoltchev, Ed., *Media Literacy, IRIS plus 2011-3* (Strasbourg, European Audiovisual Observatory, 2011), pp. 7-27.

¹⁸⁹ E.g., Nicole A. Cooke, ‘Posttruth, Truthiness, and Alternative Facts: Information Behavior and Critical Information Consumption for a New Age’, *op. cit.*, p. 219.

¹⁹⁰ PACE Resolution 2212 (2018), ‘The protection of editorial integrity’, 25 April 2018, para. 10.

5.1.3 Technological initiatives

The examples surveyed in the previous chapter are initiatives taken by tech companies themselves; they seek to identify and address specific problems or areas of concern. They can offer bit-part solutions to the bigger phenomenon of online disinformation. There is increasing political traction at the international and European levels for Internet companies/actors to exercise human rights due diligence across their activities. Pro-active measures that enhance the public's right to freedom of expression and the quality of public debate are therefore welcome, provided they are fully compliant with international and European human rights and media and communications law. When that is the case, examples such as those surveyed can be practical ways of online actors fulfilling their roles and responsibilities, as mentioned in the previous sub-section.

5.2 Identification and monitoring measures

5.2.1 Fact-checking

Fact-checking, as already mentioned, is perceived as a very important contemporary form of public-watchdog or accountability journalism. It can be undertaken by a range of actors: mainstream media and professional journalists, dedicated organisations, civil society organisations, educational organisations. Fact-checking can have a sensitizing or awareness-raising function by alerting the public to instances of inaccurate information in public debate generally and the news specifically. It can also have an educational function – if it includes an emphasis on educating the public about how to fact-check.¹⁹¹ It also has an important corrective function in respect of targeted instances of false information.

The effectiveness of fact-checking is, however, perceived as having a number of limitations. First, many fact-checking organisations are under-resourced, in terms of funding and trained/professional staff. They consequently tend to rely heavily on the engagement of voluntary efforts. It is difficult for such shoe-string organisations to compete with a swirl of mis- and dis-information from a multitude of sources, including systematic campaigns driven by large-scale, transnational, well-funded players, sometimes with strong political backing,

¹⁹¹ See Nieuwscheckers' video clip about photo verification, tweeted on 17 April 2018.

who use social media and other techniques to reinforce their messages. This is a David-and-Goliath kind of struggle, with clear inequality of arms in the communicative battle. Fact-checking is labour- and resource-intensive; it calls for scrupulous checks, which necessarily take time. And time is an elusive luxury in the relentless, 24/7 dynamics of contemporary news production and consumption. All of this leads to fact-checking lagging behind misinformation, “both in terms of overall reach and of sheer response time”.¹⁹²

Moreover, as Monroe Price has observed, in the present media ecosystem, characterized by its complexity and the multiplicity of actors and narratives, “it is hardly possible anymore to merely communicate”.¹⁹³ This is relevant for fact-checking because in order to be effective, the fact-checking must reach the section of the public that is affected by, or susceptible to, mis- and disinformation. The tendency of fact-checking to focus on individual news items or pieces of content gives it a reactionary, ad-hoc and piecemeal character. The use of social bots has been suggested as a possible strategy to ensure wider dissemination of fact-checking.¹⁹⁴ While this suggestion has a certain appeal, it could also inadvertently draw attention to, thereby leading to amplification of, the original instance of mis- or disinformation. The inadvertent amplification of the original disinformation, whether through repetition (for the purpose of refutation) or redirection of attention (again to refute it) can, then, in some cases be seen as a backfiring strategy.

The enduring impact of initial exposure to false information is also a major challenge for the effectiveness of fact-checking. Extensive literature suggests that initial information and our evaluation of information can be difficult to dislodge, even after the information has been proved to be inaccurate. Some research puts forward the thesis that “cognitive ability may not only be important in attitude formation, but also in attitude adjustment when the situation changes”.¹⁹⁵ The same research, which is based on a test that “maximally facilitated adjustment of the initial assessment”, suggests that individuals with lower cognitive ability, in particular, find it difficult to adjust their opinions and evaluations of information to

¹⁹² This observation is based on an analysis of Twitter conversations: Giovanni L. Ciampaglia (2017) ‘Fighting fake news: a role for computational social science in the fight against digital misinformation’, *Journal of Computational Social Science*, 1 (1), 147-153, at 149.

¹⁹³ Monroe E. Price, *Free Expression, Globalism and the New Strategic Communication*, *op. cit.*, p. 18.

¹⁹⁴ Giovanni L. Ciampaglia, ‘Fighting fake news’, *op. cit.*, p. 149.

¹⁹⁵ Jonas De keersmaecker and Arne Roets (2017) ‘Fake news’: Incorrect, but hard to correct. The role of cognitive ability on the impact of false information on social impressions’, *Intelligence*, 65, 107-110, p. 107.

subsequent revelations of its falsity.¹⁹⁶ Similar claims have been made concerning individuals with specific ideologies and cultural backgrounds.¹⁹⁷ While clearly relevant for fact-checking, this phenomenon of “misinformation persistence” is discussed in more detail in respect of ‘Flagging/labelling’, below.

Fact-checking may suffer – among some sections of society - from perceptions or accusations of bias or partisanship or lack of authority/legitimacy. In-house fact-checking initiatives and verification units of mainstream media or social networking services may be susceptible to such perceptions or accusations (even if they are not well-founded). Pro-active transparency measures by fact-checking organisations about their objectives, composition, working methods and sources of funding can thwart such perceptions and accusations. Collaboration with other (media) entities, adherence to the IFCN’s Code of Principles and accreditation by the IFCN, are also relevant ways of dispelling such perceptions and accusations. Reaching out to conservative (mainly) right-wing political media and organisations, engaging with them and getting them “on board”, has been suggested as an important goal for ensuring cross-spectrum support for countering mis- and dis-information and improving the quality of public debate generally.¹⁹⁸

5.2.2 Flagging / labelling / blacklisting

One of the purposes of flagging/labelling disinformation as such is to “debunk” the original message, i.e., to present “a corrective message that establishes that the prior message was misinformation”.¹⁹⁹ Corrections could be “partial” or “complete”.²⁰⁰ Examples of partial corrections include the updating of information and the provision of additional information or context which help to correct the detail or tenor of the original message. An example of complete corrections is the retraction of the original message. It appears that a “well argued, detailed debunking message” is often necessary to effectively refute the original message and

¹⁹⁶ *Ibid.*, p. 109.

¹⁹⁷ Man-pui Sally Chan, Christopher R. Jones, Kathleen Hall Jamieson and Dolores Albarraçín (2017) ‘Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation’, *Psychological Science*. 28 (11), 1531-1546, at p. 1544 (including further references).

¹⁹⁸ David Lazer, Matthew Baum, Nir Grinberg, Lisa Friedland, Kenneth Joseph, Will Hobbs and Carolina Mattsson, ‘Combating Fake News: An Agenda for Research and Action’, Final Report, Harvard Kennedy School, Shorenstein Center on Media, Politics and Public Policy, May 2017, p. 3.

¹⁹⁹ Man-pui Sally Chan *et al.*, ‘Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation’, *op. cit.*, p. 1532.

²⁰⁰ *Ibid.*

to “reduce misinformation persistence”.²⁰¹ It has been noted that “[m]essages that simply label the initial information as incorrect may therefore leave recipients unable to remember what was wrong and offer them no new model to understand the information”.²⁰² To address these shortcomings and to reinforce the effectiveness of debunking messages, Chan *et al.* have formulated three complementary recommendations:

- (a) Reduce arguments that support misinformation;
- (b) Engage audiences in scrutiny and counterarguing of misinformation, and
- (c) Introduce new information as part of the debunking message.²⁰³

Other authors have sounded a note of caution about corrective measures as they had documented “several instances of a ‘backfire effect’ in which corrections actually *increase* misperceptions among the group in question”.²⁰⁴

There are concerns that too much may be read into the fact-checking of particular news items or content. If one news item is declared inaccurate and exposed as such, what does that implicitly say about content not flagged as problematic or inaccurate? Could the practice unintentionally generate a false sense of confidence in the accuracy of non-flagged news items and content? These concerns are explored in the aptly-titled paper by Pennycook and Rand, “The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Stories Increases Perceived Accuracy of Stories Without Warnings”. One of the paper’s main conclusions is that “an implied truth effect, combined with the near impossibility of fact-checking all stories, could pose a major problem for any attempts to combat disinformation using warnings”.²⁰⁵

A potential problem with the practice of blacklisting is that errors can be made, with negative consequences for the blacklisted party. If, for instance, a news media organisation is wrongly accused of spreading disinformation, and is publicly “named and shamed”, it can sustain significant reputational damage. After all, accuracy and reliability, are the tools of the news

²⁰¹ *Ibid.*

²⁰² Man-pui Sally Chan *et al.*, ‘Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation’, *op. cit.* (with further references), p. 1532.

²⁰³ *Ibid.*, p. 1544.

²⁰⁴ (emphasis per original) Brendan Nyhan and Jason Reifler (2010) ‘When Corrections Fail: The Persistence of Political Misperceptions’, *Political Behavior*, 32 (2), 303-330, at p. 303.

²⁰⁵ Gordon Pennycook, Adam Bear, Evan T. Collins and David G. Rand, ‘The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Stories Increases Perceived Accuracy of Stories Without Warnings’, SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, 8 December 2017, at p. 4.

media's trade and if the public perceives that these values are not being upheld, trust in the media organisation will be eroded. The risk of such errors being committed can be exacerbated when the blacklisting entity does not adhere to high professional values and standards, for different reasons, like lack of resources, training, etc.

More generally in respect of monitoring “fake news” and online disinformation, the Parliamentary Assembly of the Council of Europe has recommended that member States “consider establishing a national observatory to track dissemination of disinformation, propaganda and fake news and propose adequate measures to counteract these phenomena”.²⁰⁶ This appears to be a kite-flying exercise as the idea is not explained further.

5.3 Containing or corrective measures

The contextualisation of “fake news”, for example, through the use of disputed content flags or other such signalling or warning mechanisms for the public, are welcome examples of media literacy in practice. Such measures appeal to the critical and cognitive faculties of audiences and readers. The effective response to such signals requires active engagement. In an age of information abundance and the often taken-for-granted ease of instant access to content, such signals will only be responded to by some sections of the public. The previously outlined concerns about the persistence of mis- and disinformation and the processes of attitude formation and adjustment, as well as filter bubbles, are all relevant here too. The credibility of disputed content signalling would also likely be enhanced when such determinations are made, not by any one actor (e.g. Facebook), but by a partnership or coalition of actors or organisations acting in accordance with transparent procedures.

The promotion of slow, informed and responsible sharing (“think before you click”) are also measures that respect the autonomy of the individual and are non-intrusive from a freedom of expression perspective. The closure of automated accounts could indeed be a helpful contribution to reducing online disinformation, but the techniques of identifying such accounts must not lead to generalized surveillance practices or any other practices that would unduly interfere with the rights to freedom of expression, privacy, data protection or other rights of users.

²⁰⁶ PACE Resolution 2212 (2018), ‘The protection of editorial integrity’, 25 April 2018, para. 8.7.

5.4 Regulatory or non-regulatory counter measures

All regulatory measures at the national level must comply fully with European human rights law, selected key principles of which are summarized in Chapter 3. Regulation must be prescribed by law, pursue a legitimate aim and be necessary in a democratic society and proportionate to the aim pursued. “Fake news” should not be introduced as a basis for regulation as it is overly broad and too incoherent and imprecise for regulatory purposes. A more fruitful approach would be to identify key components of “fake news” and assess the extent to which they are already covered by existing legislation. Sanctions for the dissemination of illegal content should be proportionate owing to the chilling effect they can have on freedom of expression. Regulation should take account of the roles and responsibilities of all actors, including online actors, in these matters.

As to the non-regulatory measures by private, i.e., non-State actors: credibility scoring is another example of a practical application of media literacy that offers resources to the public which it can use to make informed decisions about and assessments of content. The pros and cons of other signalling mechanisms also largely apply here. It is imperative to address advertising models and practices in appropriate, effective ways. It is widely believed that clickbait models, for instance, incentivize both the creation and the dissemination of “fake news” and disinformation.

5.5 Need for further research

This chapter opened with the caveat that academic literature on “fake news” and online disinformation is burgeoning, but limited. A standard recommendation in scholarly articles and policy documents is to stimulate and facilitate more research. As such, more and more issue-specific research centres and units are being established (see Section 4.3.1, above) and issue-specific research agendas are being adopted. A recent, high-profile collaborative research initiative between Harvard Kennedy School and Northeastern University has summarised its ‘Agenda for Research and Action’ for combating fake news as follows:

As a research community, we identified three courses of action that can be taken in the immediate future: involving more conservatives in the discussion of

misinformation in politics, collaborating more closely with journalists in order to make the truth “louder,” and developing multidisciplinary community-wide shared resources for conducting academic research on the presence and dissemination of misinformation on social media platforms.

Moving forward, we must expand the study of social and cognitive interventions that minimize the effects of misinformation on individuals and communities, as well as of how socio-technical systems such as Google, YouTube, Facebook, and Twitter currently facilitate the spread of misinformation and what internal policies might reduce those effects. More broadly, we must investigate what the necessary ingredients are for information systems that encourage a culture of truth.²⁰⁷

²⁰⁷ David Lazer *et al.*, ‘Combating Fake News: An Agenda for Research and Action’, *op. cit.*, Executive Summary.

6. Applicability in the Netherlands

A selection of different measures and initiatives against “fake news” was described in Chapter 4, and a further refined selection was briefly evaluated in Chapter 5. Before assessing the suitability of selected measures and initiatives for adoption or adaptation in the Netherlands, it is useful to first consider the extent to which, and the ways in which, the phenomenon of “fake news” has captured public attention, and fuelled concerns and debate in the Netherlands over the past few years.

“Fake news” had until very recently not really been a combustible topic in public and political debate in the Netherlands. In other countries, such as the UK, the US, France, Germany and Italy, the debate has been more intense, and for good reason – wide and well-founded suspicions of external interference in electoral and/or referendum processes. In the Netherlands, by contrast, the fear of Russian meddling in Dutch elections has been present, but (much) less palpably than in certain other countries. Moreover, such fears and concerns have been instigated in part by the experiences of other countries. Instead, the most sensitive focus of “fake news”, mis- and disinformation, has been reporting and commentary on the MH17 tragedy and the quest for facts and justice in its aftermath.

It is only since the end of 2016 that public awareness of, and concerns about, “fake news” started to ignite a national discussion. This can perhaps be explained, in part, by patterns of news consumption and levels of trust in news media in the Netherlands. In its Digital News Report of 2017, the Reuters Institute for the Study of Journalism stated that “in the Netherlands, fake news is more media hype than a real challenge, with little serious concern about its impact”. The Dutch media environment is characterized by a high level of trust in news: 51% of users trust news overall, and 62% trust the news they use. Out of the 36 researched countries, the Netherlands is in second place regarding the perceived non-existence of political and commercial influence on news. In sum, the Dutch media environment has a variety of quality news sources, in which Dutch people have high levels of trust. Due to the diverse patterns of news consumption, it is questionable to what extent the Dutch news environment serves as a fertile ground for “fake news”.

Hoax-Wijzer, a website that monitors falsehoods in online news articles, claims that around 50 Dutch and Belgian websites produce and disseminate ‘fake news’ or dubious content, often solely aimed at making profit, by generating advertising revenues through ‘clicks’. Based on its monitoring activities, it reports that since 2015 there has been a rise of inaccurate news articles in the Netherlands, also called ‘hoaxes’. Moreover, it has signalled that most of the articles are spread via social network services, which lead readers directly to hoax websites. However, not all ‘fake news’ publishers have profitmaking as their objective. According to Hoax-Wijzer, three other sources of inaccurate news can be identified in online Dutch media:

- 1) websites that publish right-wing extremist propaganda;
- 2) websites that disseminate inaccurate medical information;
- 3) websites that outline conspiracy theories.

In this context, Facebook and Twitter are used in particular, as they have a strong dominance in the inter-personal communication of users, and increasingly in other niche markets of the information and communications sphere.

6.1 Preventive or pre-emptive measures

In keeping with the consistent recommendations of research and policy documents (see Section 4.1), it is recommended that the Dutch government should continue to ensure a favourable environment for independent, quality journalism and media. In this connection, continued support for public service broadcasting and the *Stimuleringsfonds voor de Journalistiek*²⁰⁸ are very important.

On its website, the *Mediawijzer* is aptly described as the online heart of the Dutch media literacy network.²⁰⁹ The national campaign ‘Week van de Mediawijsheid’ (Week of Media Wisdom)²¹⁰ is organized annually, specifically aimed to raise awareness about media literacy amongst children and young people. The threats of “fake news” were one of the aspects of

²⁰⁸ See: <https://www.svdj.nl/>.

²⁰⁹ See: <https://www.mediawijzer.net/>.

²¹⁰ Organized by www.mediawijzer.net, see Mediawijzer, ‘Week van de Mediawijsheid 2017 – Generatie Media: samen mediawijs’, 4 September 2017, available at: <https://www.mediawijzer.net/week-van-de-mediawijsheid-2017-generatie-media-samen-mediawijs/>.

media literacy that the campaign tackled in 2017.²¹¹ As part of the campaign, a local Dutch library consciously spread “fake news” about a leak in its building on Twitter and Facebook, and thereafter spread pamphlets on how to recognize “fake news”. Similarly, ‘ Kennisnet’, a public organization for Education and ICT, sharing knowledge with the Dutch education sector, translated the infographic created by EAVI. The translated infographic sets out ten different types of “misleading news”, different motivations for publication, and different levels of influence of news. Accordingly, it emphasizes that ‘fake news’ is a broad term. The infographic will be published in a handbook which informs schools on how to integrate digital literacy in their education.

At the beginning of March, in the run-up to the municipality elections, the NOS organised an evening thematic discussion, ‘Nieuws of Nonsens’, involving a range of experts, and which it broadcast live.

These examples show the interplay between general media literacy and specific focuses within media literacy activities. They also show that different actors can contribute in different ways.

6.2 Identification and monitoring measures

Given the importance of fact-checking for public-watchdog activities in democratic society, the Netherlands appears to be lagging behind some other countries in terms of dedicated fact-checking organisations. The best-known organisation, Nieuwscheckers.nl, is a student-centric project operating from the University of Leiden: students of the master’s degree in Journalism and New Media fact-check in collaboration with ‘EenVandaag’. It was founded, and is supervised, by academics Peter Burger and Alexander Pleijter. In March 2017, Facebook started a cooperation with NU.nl and Nieuwscheckers. NU.nl and Nieuwscheckers aim to fact-check news articles on Facebook. NU.nl stated that it participates in the project because ‘fake news’ can harm the trust of people in journalism. Moreover, the website sees it as its duty to ensure that news remains ‘real’, and ‘fake news’ will not get attention.

²¹¹ *Ibid.*

Alexander Pleijter recently called for structured funding for a dedicated fact-checking organisation in the Netherlands.²¹² While structured funding for these activities is important, the critical remarks about the effectiveness of fact-checking (see above) are also valid in the Netherlands.

Some actors have argued that a journalistic quality label would enable people to distinguish real news from fake news, and suggested that the government could facilitate such a quality label, in particular the Dutch Association of Journalists (NVJ) and the Netherlands Press Council. However, the NVJ reacted that is primarily the task of the government to ensure independent journalism. Similarly, the Netherlands Press Council stated that it does not deem itself to be responsible set up such a quality label.

6.3 Containing or corrective measures

The contextualization of ‘fake news’ is one aspect of media literacy which has not yet received enough attention in the Netherlands. The attractiveness of such a measure lies in its compatibility with the right to freedom of expression as it merely provides guidance to its audience without preventing it from accessing the desired content. Unlike fact-checking measures aimed at debunking “fake news”, contextualization forces users to critically think about the nature of news articles. It is an innovative way for readers to engage in recognizing “fake news”, through the use of warnings mechanisms. As was mentioned in section 5.3, such awareness-raising measures would gain in credibility through the adoption of collaborative measures instead of merely relying on one actor. Similar partnerships to the one existing between Facebook and Nu.nl, Nieuwscheckers and Universiteit Leiden could be developed. Importantly, such measures should only serve as complementary ones as they tend to have a limited impact in terms of users being receptive to them and responding accordingly. Moreover, in order to avoid counter-productiveness or the Streisand effect, the signalling mechanisms should be well-devised.

6.4 Regulatory and non-regulatory counter measures

²¹² Alexander Pleijter, Gastcolumn – ‘In tijden van nepnieuws zijn factcheckers van maatschappelijk belang’, *de Volkskrant*, 15 April 2018. <https://www.volkskrant.nl/opinie/gastcolumn-in-tijden-van-nepnieuws-zijn-factcheckers-van-maatschappelijk-belang~a4592463/>.

In April 2017, two parliamentary members introduced a motion, in which they asked the Dutch government about the desirability and feasibility of a law which would tackle ‘fake news’. The German NetzDG was cited as an example of legislation which could be followed, as it introduces an obligation for social media companies to have clear procedures regarding illegal content and expressions on their platforms. In response to the motion, the former Minister of Security and Justice (Stef Blok) and the former Minister of the Interior and Kingdom Relations (Ronald Plasterk) explained in a letter to the parliament of July 2017 how the Netherlands deal with punishable online content. They set out that the Netherlands emphasize the importance of a common European approach. The Ministers explained that the German NetzDG imposes high fines on social media companies if they do not remove illegal content within the short reaction period set out in the law. They argued that social media companies will be more likely to remove non-illegal content, in order to minimize the risks of getting fines.

Subsequently, the former Ministers stated that the Dutch government takes the view that the online community must regulate and control itself. The former cabinet stated that national laws will be decisive when there is a conflict between law and the community standards of social media. The government has a preference for an independent legal assessment. The Minister explained that in the Netherlands, Notice-and-Take-Down (NTD) and/or Notice-and-Take-Action (NTA) procedures are applicable to illegal content. For the removal of non-illegal content, such as bullying and swearing, civil and/or criminal procedures are available. After this assessment, the former Ministers concluded that additional laws would not have added value to the current Dutch approach. Rather, enforcing new laws might even have a counter-productive effect by unnecessarily restricting freedom of expression.

Accordingly, they argued that there is no reason to follow the German model regarding illegal online content. They pointed out that European proposals to counter illegal content online will be examined critically, taking into account their negative impact on freedom of expression, as well as the consequences for users and the platforms. Nevertheless, the former Ministers stated that the Netherlands call for more transparency of social media platforms regarding their internal procedures, and that they must be open about the number of requests to remove illegal content.

The above approach is in line with what was mentioned in section 5.4, where it was submitted that preference should be given to the countering of “fake news” based on existing laws rather than on new ones. Moreover, taking into account the vagueness and ambiguity of the term “fake news”, it would be hard for a new law that is tailored to the countering of such content to fulfil the requirement of being ‘prescribed by law’, which is a *conditio sine qua non* for any legitimate interference with the right to freedom of expression (see Chapter 3).

It should be noted in passing that Kajsa Ollongren, the current Minister of Internal Affairs, discussed the dangers of online disinformation as a tool for foreign state actors in a letter to Parliament. This letter does not, however, expand on the current government’s position on regulation regarding disinformation,²¹³ so its content and the ensuing parliamentary debate will not be further discussed here.

Reliance on non-regulatory measures, such as the use of credibility scores, seems appropriate in the Netherlands. As with the ‘contextualisation’ of fake news, credibility scoring contributes to media literacy as it educates users to critically engage with news articles that might be “fake”. Relying on quality rating of users themselves, as announced by Facebook, is one technique that should not be overlooked. According to a Belgian study, conducted in 2017 by Whyte Corporate Affairs, which investigated the trust of users towards different media actors, ‘peers’ were experienced as being the most trusted actors after scientists. While automatic extrapolation of findings to the Netherlands should be avoided, considering initiatives building on users’ perceived trust would be opportune. This should, however, be complemented by other initiatives in which collaboration with professional parties should be sought, as this would enhance the overall credibility of these non-regulatory measures. As was mentioned in section 5.4, advertising models should be addressed in order to preclude the occurrence of clickbait practices. Whereas France tackles this issue through the adoption of regulatory measures, it would be less intrusive to tackle this by means of self-regulation. The focus should be for news media platforms to be more transparent with regard to the advertised content they host. Enhancing transparency is in line with what former Dutch ministers have deemed necessary to improve, when discussing the (non-)desirability of having in place Dutch regulatory measures for countering illegal content online.

²¹³ Minister Ollongren, ‘Brief aan de Voorzitter van de Tweede Kamer betreft Toezegging aan lid Buma (CDA) inzake beïnvloeding van de publieke opinie door statelijke actoren’, 13 November 2017, available at: <https://www.rijksoverheid.nl/documenten/kamerstukken/2017/11/13/kamerbrief-over-beïnvloeding-van-de-publieke-opinie-door-statelijke-actoren>.

7. Conclusion

This report recommends that the term “fake news” should be avoided as it is vague, inadequate and politically coloured. The frame of online disinformation is closer to the actual academic and policy discussions that are taking place. This recommendation is consistent with the approach taken in major studies at Council of Europe and EU levels.

The report sets out the most important reference points in the European and international human rights framework that govern online disinformation and measures to counter it. Those standards insist on strong safeguards for freedom of expression in regulatory and non-regulatory measures primarily by States, who are the addressees of those standards, but also by tech companies and the media, inasmuch as they also have responsibilities to respect human rights throughout their activities.

A great diversity of strategies and measures has already been developed to counter online disinformation. This report categorizes them in terms of their objectives: preventive or pre-emptive measures, identification and monitoring measures, containing or corrective measures, and regulatory and non-regulatory measures. The report stresses the importance of the complementary nature of such measures, and cautions against new regulatory initiatives on the basis of international human rights standards and the concerns raised over the analysed initiatives in Germany, Italy and France.

While scholarship and policy and impact analyses of the surveyed counter strategies and measures are still emerging, no hard-and-fast conclusions can be drawn about their individual or collective effectiveness. The tentative review provided reveals lots of positives but also lots of lessons for the continued development of appropriate strategies and measures in the Dutch context.

8. Bibliography

Academic works

Anderson, C.W., Downie, L. Jr. and Schudson, M. (2016) *The News Media: What Everyone Needs to Know*. New York: Oxford University Press.

Bode, L. & Vraga, E (2015) In Related News, That Was Wrong: The Correction of Misinformation Through Related Stories Functionality in Social Media, *Journal of Communication*, 65 (4): 619-638.

Chan, M. S., Jones, C. R., Jamieson, K. H. & Albarracín, D. (2017) Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological Science*. 28 (11), 1531-1546.

Chen, A. (2017) The Fake-News Fallacy. *The New Yorker*. 4 September 2017.

Ciampaglia, G. L. (2017) Fighting fake news: a role for computational social science in the fight against digital misinformation. *Journal of Computational Social Science*. 1 (1), 147-153.

Cook, J., Ecker, U. & Lewandowsky, S. (2015) Misinformation and how to correct it. *Emerging Trends in the Social and Behavioral Sciences*. 1-17.

Cooke, N. A. (2017) Posttruth, Truthiness, and Alternative Facts: Information Behavior and Critical Information Consumption for a New Age. *The Library Quarterly*. 87 (3), 211-221.

D'Ancona, M. (2017) *Post Truth: The New War on Truth and How to Fight Back*. London: Penguin.

De keersmaecker, J. & Roets, A. (2017) 'Fake news': Incorrect, but hard to correct. The role of cognitive ability on the impact of false information on social impressions. *Intelligence*. 65, 107-110.

Dunlop, B. (2017) Perfect storm: the multiple challenges facing public service news, and why tackling them is vital for democracy. Geneva: European Broadcasting Union.

Dutton, W.H. (2017) Fake news, echo chambers and filter bubbles: Underresearched and overhyped. *The Conversation*, 5 May 2017.

Dutton, W. H., Reisdorf, B. C., Dubois, E. & Blank, G. (2017) Social shaping of the politics of Internet search and networking: Moving beyond filter bubbles, echo chambers, and fake news. *SSRN Electronic Journal*.

Fletcher, Richard, Alessio Cornia, Lucas Graves, and Rasmus Kleis Nielsen. 2018. "Measuring the Reach of 'Fake News' and Online Disinformation in Europe." Oxford: Reuters Institute for the Study of Journalism.

Haigh, M., Haigh, T. & Kozak, N. I. (2017) Stopping Fake News: The work practices of peer-to-peer counter propaganda. *Journalism Studies*. 1-26.

Howard, P. & Bradshaw, S. (2017) Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation. Oxford: Computational Propaganda Project.

Jack, C. (2017) Lexicon of Lies: Terms for Problematic Information. *Data & Society*.

Kleis Nielsen, R. Cornia, A. and Kalogeropoulos, A. (2016) Challenges and opportunities for news media and journalism in an increasingly digital, mobile and social media environment, Council of Europe report DGI(2016)18.

Lazer, D., Baum, M. Grinberg, N., Friedland, L., Joseph, K., Hobbs, W. & Mattsson, C. (2017) Combating Fake News: An Agenda for Research and Action. Harvard Kennedy School, Shorenstein Center on Media, Politics and Public Policy 2.

Lewandowsky, S., Ecker, U. & Cook, J. (2017) Beyond misinformation: Understanding and coping with the "post-truth" era. *Journal of Applied Research in Memory and Cognition*. 6 (4), 353-369.

Lewandowsky, S., Ecker, U., Seifert, C.M., Schwarz, N. & Cook, J. (2012) Misinformation and Its Correction. *Psychological Science in the Public Interest*. 13 (3), 106-131.

Marwick, A. E. & Lewis, R. (2017) *Media Manipulation and Disinformation Online. Data and Society*.

McGonagle, T. (2017) “Fake news”: False fears or real concerns? *Netherlands Quarterly of Human Rights*. 35 (4), 203-209.

McGonagle, T. and Donders, Y. (eds.) (2015) *The United Nations and Freedom of Expression and Information: Critical Perspectives*. Cambridge: Cambridge University Press.

- Mendel, T., ‘The UN Special Rapporteur on freedom of opinion and expression: progressive development of international standards relating to freedom of expression’, in *ibid.*, pp. 235-268.

McGonagle, T. (2011) ‘Media Literacy: No Longer the Shrinking Violet of European Audiovisual Media Regulation?’, in S. Nikoltchev, Ed., *Media Literacy, IRIS plus 2011-3* (Strasbourg, European Audiovisual Observatory, 2011), pp. 7-27.

Newman, N. & Fletcher, R. (2017) *Bias, Bullshit and Lies: Audience Perspectives on Low Trust in the Media*. Oxford: Reuters Institute for the Study of Journalism.

Newman, N., Fletcher, R., Kalogeropoulos, A. Levy, D.A. and Nielsen, R.K. (2017) *Reuters Institute Digital News Report 2017*. Oxford: Reuters Institute for the Study of Journalism.

Nyhan, B. & Reifler, J. (2010) When Corrections Fail: The Persistence of Political Misperceptions. *Political Behavior*. 32 (2), 303-330.

Pennycook, G., & Rand D. (2017) “The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Stories Increases Perceived Accuracy of Stories Without Warnings.” SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, December 8, 2017.

Price, M.E. (2015) *Free Expression, Globalism and the New Strategic Communication*. New York: Cambridge University Press.

Schudson, M. (1995) *The Power of News*. Cambridge, Massachusetts and London, England: Harvard University Press.

Spindler, G. (2017) *Internet Intermediary Liability Reloaded – The New German Act on Responsibility of Social Networks and its (In-) Compatibility with European Law*. 8 (2017) JIPITEC 166.

Strömbäck, J., Falasca, F. & Kruikeimer, S. (2017). *The Mix of Media Use Matters: Investigating the Effects of Individual News Repertoires on Offline and Online Political Participation*. *Political Communication*.

Tandoc, E. C., Zheng Wei Lim & Ling, R. (2017) *Defining “Fake News”*. *Digital Journalism*. 1-17.

Tandoc, E. C., Ling, R., Westlund, O., Duffy, A., Goh, D. & Lim Zheng Wei. (2017) *Audiences’ acts of authentication in the age of fake news: A conceptual framework*. *New Media & Society*.

Tucker, J.A., Guess, A., Barbera, P., Vaccari, Siegel, A., Sanovich, S., Stukal, D and Nyhan, B. (2018) *Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature*.

Uberti, D. (2016). *The real history of fake news*. *Columbia Journalism Review*. 15 December 2016.

Wardle, C. & Derakhshan, H. (2017) *Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making*. Report to the Council of Europe.

Wardle, C. (2017). *Fake news. It’s complicated*. *Medium*. 16 February 2017.

Woolley, Samuel C., and Philip N. Howard. 2017. Computational Propaganda Worldwide: Executive Summary. Oxford: Oxford Internet Institute, University of Oxford.

Zuiderveen Borgesius, F., Trilling, D., Moeller, J., Bodó, B., de Vreese, C. H. & Helberger, N. (2016). Should we worry about filter bubbles? Internet Policy Review. Journal of Internet Regulation. 5 (1).

Regulatory texts

European and international treaties

European Convention on Human Rights (adopted 4 November 1950, entered into force 3 September 1953).

International Covenant on Civil and Political Rights (adopted 16 December 1966, entered into force 23 March 1976) 999 UNTS 171.

Universal Declaration of Human Rights (adopted 10 December 1948) UNGA Res 217A (III).

EU Directives

Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market, [2000] OJ L 178/1 (Directive on Electronic Commerce).

Directive (EU) 2015/1535 of the European Parliament and of the Council of 9 September 2015 laying down a procedure for the provision of information in the field of technical regulations and of rules on Information Society services.

National legislation

Germany

Act to Improve Enforcement of the Law in Social Networks (Network Enforcement Act)/Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken (Netzwerkdurchsetzungsgesetz – NetzDG), 1 September 2017, Bundesgesetzblatt, Jahrgang 2017, Teil I, Nr. 61, 7 September 2017 (entry into force: 1 October 2017).

France

The French Constitution, 4 October 1958

Loi du 29 Juillet 1881 sur la liberté de la presse

<https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006070722>

Assemblée Nationale, n. 772 - Proposition de Loi organique relative à la lutte contre les fausses informations’, 16 March 2018

<http://www.assemblee-nationale.fr/15/propositions/pion0772.asp>

Assemblée Nationale, n. 799 - ‘Proposition de Loi relative à la lutte contre les fausses informations’, 21 March 2018

<http://www.assemblee-nationale.fr/15/propositions/pion0799.asp>

European Court of Human Rights: case law

Airey v. Ireland, 9 October 1979, Series A, no. 32.

Bowman v. the United Kingdom, 19 February 1998, Reports of Judgments and Decisions 1998-I.

Chauvy and Others v. France, no. 64915/01, ECHR 2004-VI.

Colombani and Others v. France, no. 51279/99, ECHR 2002-V.

Dichand and Others v. Austria, no. 29271/95, 26 February 2002.

Dink v. Turkey, nos. 2668/07, 6102/08, 30079/08, 7072/09 and 7124/09, 14 September 2010.

Fressoz and Roire v. France [GC], no. 29183/95, ECHR 1999-I.

Handyside v. the United Kingdom, 7 December 1976, Series A, no. 24.

Informationsverein Lentia and Others v. Austria, 24 November 1993, Series A no. 276.

Jersild v. Denmark, 23 September 1994, Series A no. 298.

Lingens v. Austria, 8 July 1986, Series A no. 103.

Magyar Tartalomszolgáltatók Egyesülete and Index.hu Zrt v. Hungary, no. 22947/13, 2 February 2016.

Orlovskaya Iskra v. Russia, no. 42911/08, 21 February 2017.

Salov v. Ukraine, no. 65518/01, ECHR 2005-VIII (extracts).

Tyrer v. the United Kingdom, 25 April 1978, Series A no. 26.

X and Y v. the Netherlands, 26 March 1985, Series A no. 91.

Policy documents / miscellaneous

Council of Europe

Committee of Ministers

Declaration on freedom of political debate in the media, 12 February 2004.

Recommendation CM/Rec(2011)7 of the Committee of Ministers to member states on a new notion of media, 21 September 2011.

Recommendation CM/Rec(2012)3 of the Committee of Ministers to member States on the protection of human rights with regard to search engines, 4 April 2012.

Recommendation CM/Rec(2012)4 of the Committee of Ministers to member States on the protection of human rights with regard to social networking services, 4 April 2012.

Recommendation CM/Rec(2016)4 of the Committee of Ministers to member States on the protection of journalism and safety of journalists and other media actors, 13 April 2016.

Recommendation CM/Rec(2018)1 of the Committee of Ministers to member States on media pluralism and transparency of media ownership, 7 March 2018.

Recommendation CM/Rec(2018)2 of the Committee of Ministers to member States on the roles and responsibilities of internet intermediaries, 7 March 2018.

Parliamentary Assembly

PACE Resolution 2212 (2018), The protection of editorial integrity, 25 April 2018.

Secretary General

Secretary General of the Council of Europe, *Populism – How strong are Europe’s checks and balances?*, 4th annual report on the state of democracy, human rights and the rule of law in Europe, 2017, Chapter 2 – Freedom of Expression.

European Union

Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, *Tackling online disinformation: a European Approach*, COM(2018) 236 (final), 26 April 2018.

United Nations

Human Rights Committee

Concluding Observations: Cameroon, 4 November 1999, CCPR/C/79/Add.116.

UN High Commissioner for Human Rights

UN High Commissioner for Human Rights, Statement, 30 August 2017, available at: <http://www.unmultimedia.org/avlibrary/asset/1958/1958570/>

Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression

David Kaye, ‘Mandate of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression’ (Letter to the German authorities), Ref. no. OL DEU 1/2017, 1 June 2017, available at:

<http://www.ohchr.org/Documents/Issues/Opinion/Legislation/OL-DEU-1-2017.pdf>

David Kaye, ‘Mandate of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression’ (Letter to the Italian authorities), Ref. no. OL ITA 1/2018, 20 March 2018, available at:

<http://www.ohchr.org/Documents/Issues/Opinion/Legislation/OL-ITA-1-2018.pdf>

Miscellaneous

The United Nations (UN) Special Rapporteur on Freedom of Opinion and Expression, the Organization for Security and Co-operation in Europe (OSCE) Representative on Freedom of the Media, the Organization of American States (OAS) Special Rapporteur on Freedom of Expression and the African Commission on Human and Peoples’ Rights (ACHPR) Special Rapporteur on Freedom of Expression and Access to Information, Joint Declaration on “Fake News”, Disinformation and Propaganda, 3 March 2017.

Italy

AGCOM, Guidelines for equal access to online platforms during the election campaign for the 2018 general elections, 1 February 2018.

News articles, reports and analyses

Alexander Pleijter, Gastcolumn - 'In tijden van nepnieuws zijn factcheckers van maatschappelijk belang', *de Volkskrant*, 15 april 2018, available at: <https://www.volkskrant.nl/opinie/gastcolumn-in-tijden-van-nepnieuws-zijn-factcheckers-van-maatschappelijk-belang~a4592463/>

Alexandre Piquard and Martin Untersinger, '«Fake News » : Les pistes du texte de loi en préparation', *LeMonde.fr*, 7 March 2018, http://www.lemonde.fr/actualite-medias/article/2018/03/07/fake-news-les-pistes-du-texte-de-loi-en-preparation_5266947_3236.html

Angelique Chrisafis, 'Emmanuel Macron promises ban on fake news during elections', *The Guardian*, 3 January 2018, available at: <https://www.theguardian.com/world/2018/jan/03/emmanuel-macron-ban-fake-news-french-president>

ARTICLE 19, 'Germany: The Act to Improve Enforcement of the Law in Social Networks', Legal Analysis, August 2017, available at: <https://www.article19.org/data/files/medialibrary/38860/170901-Legal-Analysis-German-NetzDG-Act.pdf>

Benjamin Mullin and Alexios Mantzarlis, 'Facebook has a plan to fight fake news. Here's where we come in', *Poynter*, 15 December 2016, <https://www.poynter.org/news/facebook-has-plan-fight-fake-news-heres-where-we-come>

Cas Mudde, 'Why the hysteria around the 'fake news epidemic' is a distraction', *The Guardian*, 7 February 2018, available at: <https://www.theguardian.com/commentisfree/2018/feb/07/hysteria-fake-news-epidemic-distraction>.

Chay F. Hofileña, 'Fake accounts, manufactured reality on social media' *Rappler*, 9 October 2016, available at: <https://www.rappler.com/newsbreak/investigative/148347-fake-accounts-manufactured-reality-social-media>

Constellation Research, 'Sleeping Giants take on Fake News and Brands – Is your Brand Ready?', 10 January 2017, available at: <https://www.constellationr.com/sleeping-giants-take-fake-news-and-brands-your-brand-ready>

Bernd Holznagel, 'Legal Review of the Draft Law on Better Law Enforcement in Social Networks', OSCE, May 2017, p. 23, available at: <https://www.osce.org/fom/333541?download=true>

EBU, Position Paper: "Fake news" and the information disorder, Geneva, 18 April 2018, available at: https://www.ebu.ch/files/live/sites/ebu/files/Publications/Position%20papers/EBU-Position-EN-Fake_News_Disinformation-18.04.2018.pdf

EDRI, Recommendations on the German bill "Improving Law Enforcement on Social Networks" (NetzDG)', 20 June 2017, available at: https://edri.org/files/consultations/tris_netzdg_edricontribution_20170620.pdf

Emma Jane Kirby, The city getting rich from fake news, BBC News, 6 December 2016, available at: <http://www.bbc.com/news/magazine-38168281>.

Ernesto Apa & Filippo Frigerio, 'Italy: Self-regulatory guidelines for online platforms for next general elections published by Italian Communication Authority', *IRIS – Legal Observations of the European Audiovisual Observatory* 2018-3/24, available at: <https://merlin.obs.coe.int/iris/2018/3/article24.en.html>.

Ernesto Apa & Marco Bassini, 'Italy: Legislative proposal on fake news', *IRIS – Legal Observations of the European Audiovisual Observatory*, 2017-5/27, available at: <https://merlin.obs.coe.int/iris/2017/5/article27.en.html>

European Commission, Final results of the Eurobarometer on fake news and online disinformation, Consultation results, 12 March 2018, available at: <https://ec.europa.eu/digital-single-market/en/news/final-results-eurobarometer-fake-news-and-online-disinformation>

European Commission, ‘Notification Detail – Act improving law enforcement on Social Networks’, 28 June 2017, available at: <http://ec.europa.eu/growth/tools-databases/tris/en/search/?trisaction=search.detail&year=2017&num=127>

Facebook Newsroom, ‘Blocking ads from pages that repeatedly share false news’, 28 August 2017, available at: <https://newsroom.fb.com/news/2017/08/blocking-ads-from-pages-that-repeatedly-share-false-news/>

First Draft, ‘Launching our new Chrome extension NewsCheck’, 12 January 2017, available at: <https://firstdraftnews.org/launching-new-chrome-extension-newscheck/>

Google Blog Post, ‘How we fought bad ads, sites and scammers in 2016’, 25 January 2017, available at: <https://www.blog.google/topics/ads/how-we-fought-bad-ads-sites-and-scammers-2016/>

Google Blog Post, ‘Our latest quality improvements for Search’ 25 April 2017, available at: <https://blog.google/products/search/our-latest-quality-improvements-search/>

Human Rights Watch, Germany: Flawed Social Media Law NetzDG is Wrong Response to Online Abuse, 14 February 2018, available at: <https://www.hrw.org/news/2018/02/14/germany-flawed-social-media-law>

Joel Simon, ‘Trump Is Damaging Press Freedom in the U.S. and Abroad’, *The New York Times*, 25 February 2017, available at: <https://www.nytimes.com/2017/02/25/opinion/trump-is-damaging-press-freedom-in-the-us-and-abroad.html?mcubz=3>

Katharine Viner, ‘How technology disrupted the truth’, *The Guardian*, 12 July 2016, available at: <https://www.theguardian.com/media/2016/jul/12/how-technology-disrupted-the-truth>.

Knack.be, 'Google geeft Roularta 400.000 euro om fake news te detecteren', 18 December 2017, available at: <http://datanews.knack.be/ict/google-geeft-roularta-400-000-euro-om-fake-news-te-detecteren/article-normal-940857.html>

Marc Rees, 'Exclusif, téléchargez la future loi contre les Fake News', *Nextinpact*, 7 March 2018, available at: <https://www.nextinpact.com/news/106262-exclusif-telecharger-future-loi-contre-fakes-news.htm>

Natasha Lomas, 'Facebook culls tens of thousands of fake accounts ahead of UK election', *TechCrunch*, 8 May 2017, available at: <https://techcrunch.com/2017/05/08/facebook-culls-tens-of-thousands-of-fake-accounts-ahead-of-uk-election/>

NOS, 'Media trekken kort geding tegen EU om nepnieuws in', 13 March 2018, available at: <https://nos.nl/artikel/2222185-media-trekken-kort-geding-tegen-eu-om-nepnieuws-in.html>.

NOS, 'Ook Ollongren wil nu af van Europese waakhond nepnieuws', 9 March 2018, available at: <https://nos.nl/artikel/2221353-ook-ollongren-wil-nu-af-van-europese-waakhond-nepnieuws.html>.

NOS, 'Tweede Kamer stemt voor opheffing nepnieuwsbureau EU', 6 March 2018, available at: <https://nos.nl/artikel/2220926-tweede-kamer-stemt-voor-opheffing-nepnieuwsbureau-eu.html>.

NOS, 'GeenStijl, TPO en Persgroep slepen EU voor de rechter', 20 February 2018, available at: <https://nos.nl/artikel/2218452-geenstijl-tpo-en-persgroep-slepen-eu-voor-de-rechter.html>.

Open VLD, 'Sven Gatz lanceert het 4^e Burgerkabinet', 31 January 2018, available at: <http://www.openvld.be/?type=nieuws&id=1&pageid=96598>

Ricardo Bilton, 'Reddit's /r/worldnews community used a series of nudges to push users to fact-check suspicious news', *Nieman Lab*, 2 February 2017, available at <http://www.niemanlab.org/2017/02/reddits-rworldnews-community-used-a-series-of-nudges-to-push-users-to-fact-check-suspicious-news/>

RT France, 'Le gouvernement a transmis sa loi contre les 'Fake News' aux députés LREM... mais elle a déjà fuité', 7 March 2018, available at: <https://francais.rt.com/france/48558-projet-loi-gouvernement-contre-fake-news-communique-lrem-deja-fuite>

RTL Info, 'Des Belges s'organisent pour lutter à leur manière contre les 'fake news' qui influencent l'opinion publique: voici leur stratégie', 13 March 2017, available at: <https://www.rtl.be/info/vous/temoignages/des-belges-s-organisent-pour-lutter-a-leur-maniere-contre-les-fake-news-qui-influencent-l-opinion-publique-voici-leur-strategie-899208.aspx>

Safia Samee Ali and Associated Press, 'Facebook shuts down 30 000 fake Accounts in France ahead of Presidential Elections', *NBC News*, 15 April 2017, <https://www.nbcnews.com/news/world/facebook-shuts-down-30-000-fake-accounts-france-ahead-presidential-n746821>

Santa Clara University, 'Backstory: Trust Indicators Journey', available at: <https://www.scu.edu/ethics/focus-areas/journalism-ethics/programs/the-trust-project/backstory-trust-indicators-journey/>

Sheera Frenkel and Sapna Maheshwari, 'Facebook to Let Users Rank Credibility of News', *The New York Times*, 19 January 2018, available at: <https://www.nytimes.com/2018/01/19/technology/facebook-news-feed.html>

Sleeping Giants, 'How to be a giant', 27 November 2016, available at: https://twitter.com/slpng_giants/status/802915214563807232?ref_src=twsrc%5Etfw&ref_url=http%3A%2F%2Fwww.slate.com%2Farticles%2Fnews_and_politics%2Fpolitics%2F2016%2F12%2Fsleeping_giants_campaign_against_breitbart.html&tfw_site=slate

Staff and agencies, 'Germany approves plans to fine social media firms up to €50m', *The Guardian*, 30 June 2017, available at: <https://www.theguardian.com/media/2017/jun/30/germany-approves-plans-to-fine-social-media-firms-up-to-50m>

Tessa Lyons, ‘Replacing Disputed Flags with Related Articles’, *Facebook Newsroom*, 20 December 2017, available at: <https://newsroom.fb.com/news/2017/12/news-feed-fyi-updates-in-our-fight-against-misinformation/>

Xiao Mina, ‘Building Technical Standards for Credibility’, 3 March 2017, available at: <https://misinfocon.com/building-technical-standards-for-credibility-59ef9ee4ab73>

Websites cited

Accountable journalism – monitoring media ethics across the globe
<https://accountablejournalism.org/>

Burgerkabinet Belgium
<http://www.burgerkabinet.be/>

Conseil Supérieur de l’éducation aux medias, ‘quinzaine de l’éducation aux médias’
www.csem.be/outils/operations/la_quinzaine_de_leducation_aux_medias

CrossCheck
<https://crosscheck.firstdraftnews.org/france-en/>

Data & Society
<https://datasociety.net/about/>

Data & Society Disinformation Action Lab (DAL)
<https://datasociety.net/research/disinformation-action-lab/>

EAVI ‘Beyond Fake News’ infographic
<https://eavi.eu/beyond-fake-news-10-types-misleading-info/>

Ethical Journalism Network
<http://ethicaljournalismnetwork.org/tag/fake-news>

EU vs Disinformation

<https://euvsdisinfo.eu/>

Face2Face: Real-time Face Capture and Reenactment of RGB Videos

<http://niessnerlab.org/projects/thies2016face.html>

Faktisk

<https://www.faktisk.no/>

First Draft

<https://firstdraftnews.org/about/>

Google News Lab

<https://newsinitiative.withgoogle.com/training/>

High Level Expert Group on Fake News and Online Disinformation final report

<https://ec.europa.eu/digital-single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation>

IFLA infographic 'How to Spot Fake News'

<https://www.ifla.org/publications/node/11174>

Information Disorder report

<https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-research/168076277c>

InVID innovation

<http://www.invid-project.eu/tools-and-services/invid-verification-plugin/>

Journalism Trust Initiative

<https://rsf.org/en/news/rsf-and-its-partners-unveil-journalism-trust-initiative-combat-disinformation>

The LSE Truth, Trust and Technology (T3) Commission

<http://www.lse.ac.uk/media-and-communications/truth-trust-and-technology-commission>

Mediawijs, 'Dossier : Nieuws- en Informatiegeletterdheid', 2017,

<https://mediawijs.be/dossiers/dossier-nieuws-en-informatiegeletterdheid>

Mediawijzer

<https://www.mediawijzer.net/>

Nieuws in de Klas

<https://www.nieuwsindeklas.be/contact/>

Poynter, 'Application process for signatories of the Fact-Checkers' Code of Principles'

<https://docs.google.com/document/d/1TVH6Xduaz8lYxvnRfMzi85PMTxCseNoUQ-gcdqIsnoI/edit>

Poynter, 'International Fact-Checking Network fact-checkers' code of principles'

<https://www.poynter.org/international-fact-checking-network-fact-checkers-code-principles>

Stimuleringsfonds voor de journalistiek

<https://www.svdj.nl/>

Trust Project

<https://thetrustproject.org/>