Counter-disinformation around the World: Comparing State Actions

Giovanni De Gregorio[*] and Roxana Radu[**]

Paper prepared for the 2019 Annual GigaNet Symposium (25 November, Berlin)

Keywords: Disinformation; Regulation; Democracy; Authoritarianism; Freedom of Expression

1. Introduction

Disinformation is nothing new. The primary difference since the advent of new digital media is how much of it there is, how fast it spreads and how far it reaches. In the last couple of years, the online distribution of false information has raised serious concerns worldwide.[1] The risk for democracy, the threat for fundamental rights and the role of traditional media outlets are only some of the primary topics addressed in the aftermath of events like the 2016 Brexit referendum or the last US presidential election.[2]

The web became one of the primary sources of information and knowledge for the majority of those with Internet access. Although traditional channels of information such as television and newspapers still play an important role in disseminating information, users increasingly rely on social media to get their news. The possibility to produce, distribute and access information directly from personal devices makes online content a powerful tool to influence public opinion, and, consequently, the whole society. Moreover, the increase of information sources has led, on the one hand, to the increase of the possibility to access information online. However, on the other hand, the complex assessment of the vast amount of information does not allow users to select the most reliable sources so that this situation mitigates the positive effect deriving from the increase of online media pluralism.[3]

---

[*] PhD Candidate in Public Law at University of Milano-Bicocca.

[**] Postdoctoral researcher, Programme in Comparative Media Law and Policy at the Centre for Socio-Legal Studies, University of Oxford.

[1] Bertin Martens and Others, 'The Digital Transformation of News Media and the Rise of Disinformation and Fake News' (2018) JRC Digital Economy Working Paper no. 2 <https://ec.europa.eu/jrc/sites/jrcsh/files/jrc111529.pdf>.

[2] Pew Research, 'About 6-in-10 Americans get news from social media' (2016) <http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/pj_2016-05-26_social-media-and-news_0-01/>.

[3] Cass Sunstein, *#Republic: Divided Democracy in the Age of Social Media* (Princeton University Press 2017).

Against this scenario, States have adopted various countermeasures around the world, ranging from creating a legal basis for regulating disinformation to acting without a legal basis in shutting down internet access or access to particular services on discretionary grounds.[4] Between these two categories stand the countries that had a public debate around the need to regulate but decided not to intervene, which provide a useful lens for understanding the full range of options in dealing with falsehood and public alarming. Other States have criminalised the spread of disinformation, either by introducing new laws (e.g. Singapore and Russia) or by expanding the scope of existing legislation (e.g. Saudi Arabia). Others still have adopted the last resort remedy consisting of shutting down social media (e.g. Sri Lanka and India). States like Australia and the UK have decided not to regulate the spread of disinformation, promoting debates instead, including via reports on media literacy and task forces to define national strategies to tackle this phenomenon.

Within this framework, our research explores challenges in developing countermeasures to disinformation proposed and implemented by States in the form of new legislation. The purpose of this study is to define some of the primary trends in regulatory countermeasures that governments have adopted to address online falsehood. Methodologically, our analysis relies on a self-constructed dataset that examines original legislative texts passed between 2016-2019 by states worldwide in order to counter disinformation. As of June 2019, thirty-seven countries across the globe had adopted legislation in this field or had held a public debate about the possibility of introducing such a bill. Among these, a few have interpreted existing regulation in light of the challenges posed by disinformation, thus extending the scope of previous laws. Thirteen countries had passed new laws: Belarus, Cambodia, Chile, China, Egypt, France, Germany, Israel, Kenya, Malaysia, Myanmar, Singapore, Vietnam.

After an initial mapping of what the new legislation on disinformation consists in, we proceed with examining various approaches to regulation in democratic and authoritarian regimes by using a graph made of two axes (Figure 1). The horizontal axis represents the continuum between authoritarianism and democracy. The vertical axis focuses on the regulatory approach, ranging from no action (soft law) to introducing legislation (hard law). In the latter case, the primary criteria for different degrees of regulation follow a proportionality approach based on sanctioning mechanisms and the scope of application (lowest score for regulation indirectly affecting disinformation to highest score for criminal sanctions at a general level for spreading false news). More specifically, the analysis focuses on whether the regulation in question covers a specific sector or applies generally, targets individuals or online platforms, and provides criminal or other forms of sanctions for failure to comply with legal provisions.

[4] Olga Robinson, Alistair Coleman and Shayan Sardarizadeh, 'A Report on Antidisinformation Initiative' (2019) <https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/08/A-Report-of-Anti-Disinformation-Initiatives>; Giovanni De Gregorio and Elena Perotti, 'Tackling Disinformation around the World' (2019) <https://www.wan-ifra.org/reports/2019/05/03/public-affairs-media-policy-briefing-tackling-disinformation-around-the-world>.

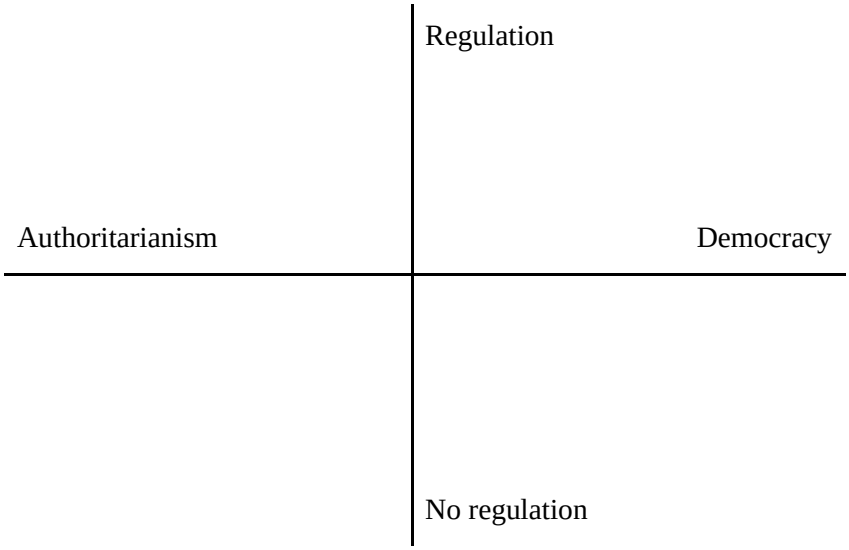

Figure 1. Conceptual outline of the relationship between type of regime and form of regulation

For the purposes of this paper, our analysis focuses on bills, acts, laws passed by national Parliaments to address online falsehood and disinformation. In other words, we only examine the upper side of the graph looking at some examples of regulation introduced by democratic and authoritarian regimes. Furthermore, we only consider new legislation without taking into account amendments to previous regulation or courts' decisions extensively interpreting existing laws to include disinformation under their scope of application.

The remainder of this paper is divided as follows. The first part introduces the topic of disinformation from a regulatory perspective, outlining why States around the world have approached the issue of disinformation in divergent ways. The second part examines specific, newly-introduced legislation, comparing approaches to the challenges raised by online disinformation. The third part provides concluding remarks, discussing main developments and trends in the fight against disinformation.

2. Understanding the Legal Fragmentation of the Disinformation Arena

Before analysing new regulatory strategies, it is worth delineating the boundaries of the arena in which States are fighting to tackle disinformation, a worldwide phenomenon exceeding their boundaries. It would not be enough to simply describe national legislation or, more generally, regulatory attempts in a comparative perspective without a preliminary examination of the key actors involved in the (digital) disinformation arena and their dominant interests.

Regulating disinformation is more intricate than it might look like at first glance.[5] Online disinformation is a cross-border issue, which requires implementation by private intermediaries both

[5] Chris Marsden and Trisha Meyer, 'Regulating Disinformation with Artificial Intelligence' (2019) European Parliamentary Research Service <https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624279/

locally and on a global scale, often resorting to artificial intelligence tools, all while preserving trust in the digital environment. In light of these complexities, the approaches to this issue are highly fragmented around the world. Some laws extend their scope to natural persons and/or legal entities (e.g. social media) and provide different forms of sanctions for failure to comply with removal obligations or the spread of alleged disinformation content. This legal fragmentation does not occur by chance; it is the result of different values and various configurations of actors involved in the fight against disinformation.

When States address online disinformation, they reflect upon balancing the interests at stake, including the role(s) assigned to information intermediaries. Indeed, tackling disinformation requires public actors to ponder whether online speech needs to be protected and if so, how; it also pushes them to think about the pursuit of other (legitimate) interests, such as public safety. Most countries around the world already have provisions against the spread of false information and manipulation in their media regulations. Whereas the protection of freedom of expression and information is self-explanatory in many contexts, it can clash with many other –legitimate or illegitimate – interests. In the disinformation arena, it does do in a visible manner.

Clues for potentially mitigating such normative conflicts can be found in the many protections afforded to freedom of expression. The roots of the right to freedom of expression show how Western democracies have been firmly influenced by a liberal approach. Already in the seventeenth century, Milton argued that the possibility to express opinions and ideas should not be restricted since the truth only prevails when freedom of expression is not threatened.[6] It is worth recalling how Milton compares the truth to a streaming fountain whose water should not be polluted by public actors' interferences. Only the free flow of information can save men from prejudice and allow them to reach knowledge and awareness. By the same token, Mill shared the same liberal approach concerning the possibility to regulate speech and contain the spread of false information. According to Mill, even falsehood could contribute to reaching the truth.[7] Indeed, censoring false opinions would not only undermine the comparison between different views but, broadly, would lead to the dogmatisation of the current truth.[8]

The scope of these liberal ideas against the interference of public actors to preserve freedom of expression was reconfirmed in the twentieth century, in Justice Holmes' dissenting opinion in *Abrams v United States* of 1919.[9] Justice Holmes argued that, although men try to support their positions by

EPRS_STU(2019)624279_EN.pdf>.

[6] John Milton, *Aeropagitica* (1644). Milton argued: 'So Truth be in the field, we do injuriously, by licensing and prohibiting, to misdoubt her strength. Let her and Falsehood grapple; who ever knew Truth put to the worse, in a free and open encounter?'

[7] John S. Mill, *On Liberty* (1859). 'First, if any opinion is compelled to silence, that opinion may, for aught we can certainly know, be true. To deny this is to assume our own infallibility'.

[8] Ibid, 'Thirdly, even if the received opinion be not only true, but the whole truth; unless it is suffered to be, and actually is, vigorously and earnestly contested, it will, by most of those who receive it, be held in the manner of a prejudice, with little comprehension or feeling of its rational grounds. And not only this, but, fourthly, the meaning of the doctrine itself will be in danger of being lost, or enfeebled, and deprived of its vital effect on the character and conduct: the dogma becoming a mere formal profession, inefficacious for good, but cumbering the ground, and preventing the growth of any real and heartfelt conviction, from reason or personal experience'.

[9] *Abrams v United States* [1919] 250 U.S. 616. 'Persecution for the expression of opinions seems to me perfectly

criticising opposing ideas, they must not be persuaded that their opinions are certain. Only the free exchange of ideas can confirm the accuracy of each position creating a 'free marketplace of ideas'.[10]

However, if these considerations show that there are reasons to protect false expressions and thus, to limit attempts made by States to regulate speech, it is also necessary to observe that not all States follow this. It is sufficient to cross the Atlantic to understand how this general trust in a vertical and negative paradigm of free speech is not shared worldwide by other democracies.[11] Unlike in the US, the paradigm of freedom of expression in Europe is subject to careful balancing between other fundamental rights and (conflicting) legitimate interests.[12] Otherwise, as the Charter of Fundamental Rights of the European Union states, granting absolute protection to one right could lead to the destruction of other fundamental rights undermining *de facto* their constitutional relevance.[13] If we move away from liberal approaches, other rationales apply: in authoritarian contexts, the strict control of information is key to limiting the threats posed to the regime, whether expressed in mild forms of censorship or in repressive measures. Consequently, the state has final authority over the flow of information, scrutinizing both public and private platforms that might enable free speech or disinformation.

We now turn to the constitutional asymmetries between democratic and authoritarian States in dealing with disinformation. As a matter of fact, authoritarian and totalitarian regimes are characterised by the predominance of a central authority. While in totalitarian regimes the central authority exercises a total power without tolerating any form of disobedience, authoritarianism aims

---

logical. If you have no doubt of your premises or your power and want a certain result with all your heart you naturally express your wishes in law and sweep away all opposition [...] But when men have realized that time has upset many fighting faiths, they may come to believe even more than they believe the very foundations of their own conduct that the ultimate good desired is better reached by free trade in ideas. […] The best test of truth is the power of the thought to get itself accepted in the competition of the market, and that truth is the only ground upon which their wishes safely can be carried out'.

[10] This expression was coined for the first time by Justice Douglas in *United States v Rumely*. *United States v Rumely* [1953] 345 U.S. 41. 'Of necessity I come then to the constitutional questions. Respondent represents a segment of the American press. Some may like what his group publishes; others may disapprove. These tracts may be the essence of wisdom to some; to others their point of view and philosophy may be anathema. To some ears their words may be harsh and repulsive; to others they may carry the hope of the future. We have here a publisher who through books and pamphlets seeks to reach the minds and hearts of the American people. He is different in some respects from other publishers. But the differences are minor. Like the publishers of newspapers, magazines, or books, this publisher bids for the minds of men in the market place of ideas'. See Oreste Pollicino, 'Fake news, Internet and Metaphors' (2017) 1(1) Rivista di diritto dei media 23; Daniel E. Ho and Frederik Schauer, 'Testing the Marketplace of Ideas' (2015) 90 New York University Law Review 1161; Eugene Volokh, 'In Defense of the Market Place of Ideas / Search for Truth as a Theory of Free Speech Protection' (2011) 97(3) Virginia Law Review 591; Alvin I. Goldman and James C. Cox, *Speech, Truth, and the Free Market for Ideas* (Cambridge University Press 1996); Ronald Coase, 'Markets for Goods and Market for Ideas' (1974) 64(2) American Economic Review 1974.

[11] Oreste Pollicino and Marco Bassini, 'Free Speech, Defamation and the Limits to Freedom of Expression in the EU: A Comparative Analysis', in Andrej Savin and Jan Trzaskowski (eds), Research Handbook on EU Internet Law (Edward Elgar, 2014); Vincenzo Zeno-Zencovich, *Freedom of Expression. A Critical and Comparative Analysis* (Routledge, 2008).

[12] Charter of Fundamental Rights of the European Union [2012] OJ C326/12, Art 52. European Convention on Human Rights [1950], Art 10(2).

[13] Charter, Art 54; Convention, Art 17.

to avoid constitutional obligations and principles such as the rule of law.[14] In the absence of any safeguard and tolerance for pluralism, regulating disinformation is not a matter of ensuring freedom of expression any longer.[15] Rather, it is an opportunity for authoritarian and totalitarian regimes to foster their legal narrative around legitimate interests such as public security to dismantle even good speech by imposing high censoring mechanism.[16]

While authoritarian regimes aim to suppress or control the degree of pluralism to avoid any interference with the central authority, democratic States are open environments for pluralism. Constitutional guarantees could be absent or neglected by autocrats, but that would not be the case for democracies: the respect of fundamental rights and freedoms, especially freedom of expression, is at the core of the entire democratic system.[17] Therefore, one of the primary challenges for democratic States when regulating disinformation is pursuing the protection of their own legitimate interests while taking into consideration other constitutional interests.

The aforementioned situation also affects the regulation of online disinformation. Indeed, the digital environment amplifies the situation due to the peculiarities of the medium of dissemination. The role of online platforms, including social media, is important in analysing the spread of false and misleading information, whether as a function of a message 'becoming viral' or as an algorithmic system pushing certain messages to the top and/or promoting certain engagement features. As observed by Balkin, in the information society, freedom of expression is like a triangle.[18] The regulation of speech does not involve any longer just the States and the speaker, but also multiple players outside the control of the State, such as social media companies. Unlike traditional media outlets, social media usually perform content moderation activities implementing automated systems which can decide in a heartbeat whether to maintain or delete the vast amount of online content globally.

Looking at how social media and search engines amplify the reach and visibility of online messages, it is possible to analyse how democratic and authoritarian States react to the challenges raised by these private actors. Due to the asymmetries between authoritarianism and democracy, national approaches strongly diverge. Indeed, when regulating online intermediaries, democratic States need to strike a fair balance between different rights and interests at stake like the freedom to conduct business of online platforms or users' freedom of expression. Unlike authoritarian regimes, democratic States cannot disregard the protection of fundamental rights and freedom. As observed by

[14] Tom Ginsburg and Alberto Simpser (eds), *Constitutions in Authoritarian Regimes* (Cambridge University Press 2014).

[15] Authoritarian countries do not deny constitutional principles and limits but manipulate them as an instrument to pursue political purposes transforming political constitutions into façade. Giovanni Sartori, 'Constitutionalism: A Preliminary Discussion' (1962) 56(4) The American Political Science Review 853.

[16] Justin Clark and Others, 'The Shifting Landscape of Global Internet Censorship' (2017) Berkman Klein Center for Internet & Society Research Publication <https://dash.harvard.edu/handle/1/33084425>.

[17] This consideration shows why fundamental rights and democracy are substantially intertwined. Because of this substantive relationship, fundamental rights cannot easily be exploited to pursue particular political ends. Susan Marks, *The Riddle of All Constitutions: International Law, Democracy, and the Critique of Ideology* (Oxford University Press 2004).

[18] Jack Balkin, 'Free Speech is a Triangle' (2018) 118 Columbia Law Review 2011.

the US Supreme Court in *Packingham v North Carolina*,[19] 'it is cyberspace – the "vast democratic forums of the Internet" in general, and social media in particular'.[20] Therefore, social media would enjoy a safe constitutional area of protection under the First Amendment which in the last twenty years, has constituted a fundamental ban on any attempt to regulate speech or bind online platforms to comply with the new obligations concerning online content.[21] On the contrary, authoritarian regimes consider this 'democratic forums' as a risk that can undermine the stability of the central authority. It is not by chance that Internet shutdowns and other intrusive forms of digital censorship such as social media blocking have spread especially in authoritarian regimes implementing these practices also to address the issue of disinformation.[22]

Against this background, the next section describes a set of regulatory strategies recently adopted by States to tackle disinformation, highlighting similarities and differences in new legislation passed on several continents.

3. Regulatory Strategies against Disinformation around the World

Since 2016, thirteen countries around the world have passed new legislation to deal with phenomena related to online disinformation, an umbrella term encompassing, as per national definitions, rumours, fake news, disinformation, falsehood or inaccurate and misleading information. Of the total number of countries legislating on the issue, we selected representative cases belonging to the tri-partite Freedom House categorization of countries, distinguishing between free, partly free and not free. Two countries have been included in each category, and an effort was made to represent as many continents as possible. While this is not an exhaustive analysis, each of the cases included here bring forward, in a succinct manner, the context for the adoption of the new law, revealing different concerns and different values.

Based on the peculiarities of the disinformation arena as described in section 2, the analysis focuses on three features to understand whether the law: 1) covers a specific sector and/or applies generally; 2) targets individuals and/or online platforms for disseminating false information; 3) establishes criminal and/or other sanctions for failure to comply with legal provisions. Important for our comparative effort, these criteria are not based on regime characteristics (e.g. respect/disregard for rule of law) which usually belong to democracies or non-democracies. On the contrary, these criteria focus on objective elements of a regulation, specifically: the scope of application and the type of sanctions for failure to comply with legal obligations.

3.1 Free States

France

[19] *Packingham v North Carolina* [2017] 582 U.S. ___.
[20] Ibid.
[21] See, for instance, *Reno v ACLU* 521 U.S. 844 (1997).
[22] Access Now, 'The State of Internet Shutdowns around the World' (2018) <https://www.accessnow.org/the-state-of-internet-shutdowns-in-2018/>.

In March 2018, the French National Assembly voted two bills, for a framework and an ordinary law, presented by president Emmanuel Macron's party, *La République en Marche*, to prevent foreign propaganda and disinformation with a specific focus on electoral periods.[23] The two proposed laws have been highly questioned, so that the Senate rejected the two proposals before finally approving the amended bills.[24] The acts were finally approved in November 2018,[25] and promulgated, after, however, passing a preliminary constitutionality review.[26]

The primary goal of this regulation is to fight the massive and rapid dissemination of false news spread by broadcasters and online providers and mitigate the challenges coming from the interference of third States. Rather than punishing the authors of misinformation, the French law aims to prevent its dissemination through social media or by broadcast. This is also reflected in the name of the legislation, focused not just on 'disinformation', but also on the conduct of 'information manipulation'.[27] More specifically, this regulation, firstly, amends the Electoral code by introducing transparency obligations for platforms and an emergency procedure for removing false information ('*action judiciaire en référée*').[28] Then, the law modifies the provisions of the *Loi relative à la liberté de communication* establishing new powers of the French audiovisual regulatory authority (*Conseil supérieur de l'audiovisuel*).[29] Moreover, the French law also introduces duties of cooperation for online platforms,[30] and measures to promote media literacy.[31]

Concerning transparency obligations, the law requires online platforms, *opérateur de plateforme en ligne*,[32] to disclose information to the public.[33] In particular, firstly, online platforms are required to deliver user with fair, clear and transparent information: 1) about the identity of the natural person or the company information (e.g. registered office), which pays online platforms to promote news

[23] Proposition de loi organique relative à la lutte contre les fausses informations, n° 772, 21 March 2018. <http://www.assemblee-nationale.fr/15/propositions/pion0772.asp>; Proposition de loi relative à la lutte contre les fausses informations, n° 799, 21 March 2018 <http://www.assemblee-nationale.fr/15/propositions/pion0799.asp>.

[24] In July 2018, the Senate rejected both bills on grounds that they risked undermining freedom of expression. Due to the extremely vague definition of fake news. Micheal-Ross Fiorentino, 'France passes controversial 'fake news' law' (*Euronews*, 22 November 2018) <https://www.euronews.com/2018/11/22/france-passes-controversial-fake-news-law>.

[25] Loi organique n° 2018-1201 du 22 décembre 2018 relative à la lutte contre la manipulation de l'information <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000037847556&dateTexte=&categorieLien=id>; Loi n° 2018-1202 du 22 décembre 2018 relative à la lutte contre la manipulation de l'information <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000037847559&categorieLien=id>.

[26] See Décision n° 2018-773 DC du 20 décembre 2018 <https://www.conseil-constitutionnel.fr/decision/2018/2018773 DC.htm>; Décision n° 2018-774 DC du 20 décembre 2018 <https://www.conseil-constitutionnel.fr/decision/2018/ 2018774DC.htm>.

[27] The government has clarified that journalistic works do not fall under the scope of application of this law since the aim is to tackle deliberate attempts to manipulate information <https://www.gouvernement.fr/action/lutte-contre-la-manipulation-de-l-information>.

[28] Loi n° 2018-1202, Title I.

[29] Ibid, Title II.

[30] Ibid, Title III.

[31] Ibid, Title IV.

[32] Code de la consommation, Article L111-7.

[33] Loi n° 2018-1202, Art 1(2).

content related to a debate of general interest; the use of his or her personal data in the context of the promotion of information content related to a debate of general interest; 3) the amount of remuneration received in return for the promotion of such information content when the amount exceeds a certain threshold. This information is aggregated in a register made available to the public by electronic means, in an open format, and regularly updated during the election period defined by this law. The French law also provides that failure to comply with the aforementioned transparency obligations can be punished with a year of imprisonment and the fine of 75000 euro.[34] Moreover, legal entities can also be subject to the sanction established by the French criminal code.[35]

Monitoring activity over these duties is tasked to the audiovisual regulatory authority, which may prevent, suspend or interrupt the broadcasting services controlled by or under the influence of a foreign State.[36] Specifically, the supervisory authority can refuse the conclusion of an agreement for the purpose of broadcasting a radio or television service if the broadcasting of that service involves a 'serious risk of violating human dignity, the freedom and property of others, the pluralistic nature of the expression of currents of thought and opinion, the protection of children and adolescents, the protection of public order, the needs of national defense or the fundamental interests of the Nation, including the proper functioning of its institutions.[37] Furthermore, within the electoral period, the audiovisual supervisory authority can also order the suspension of the broadcasting by any means of electronic communication until the end voting operations, if it finds that the service based on an agreement concluded with a legal person controlled by a foreign State or placed under the influence of that State, deliberately disseminates false information likely to affect the ballot.[38]

Besides, in the three months preceding a national election, without prejudice to the compensation for the damage suffered, the public prosecutor's office, any candidate, any party or political group or of any person having an interest in acting, who detect false information would be entitled to request a judicial decision within 48 hours to order access and hosting providers to adopt proportional and necessary measures to cease the dissemination of the content at stake.[39] False news is more precisely defined, due to the introduction of a three-step test for the identification of the illicit content. The final text establishes that the judge will identify a piece of information as false if 1) the news is manifestly false, 2) it is distributed massively and through artificial means, and 3) it is aimed at interfering with public peace or truthfulness of the electoral process.[40]

---

[34] Ibid, Art 1(1).
[35] Ibid.
[36] Ibid, Arts 5-6.
[37] Ibid, Art 5.
[38] Ibid, Art 6.
[39] Ibid, Art 1(2).
[40] In May 2019, the Paris Tribunal de grande instance delivered its first judgement in a case concerning the request of removal of a tweet posted by the Minister of Interior for alleged dissemination of false facts. The Court found that the conditions were not met. First, the Minister's tweet concerned an event that had occurred; second, the content was not sponsored with the aim to increase its dissemination requirement also through the use of artificial systems like bots; third, despite the exaggerated language, several newspapers and the Minister himself had clarified the facts. Tribunal de grande instance de Paris, 17 May 2019, n° 19/53935.

Germany

In June 2017, Germany adopted the *Netzdurchsetzunggesetz*, known also as Network Enforcement Act (NetzDG) which entered into force on 1 January 2018.[41] The aim of the law is to regulate the procedure of handling complaints regarding unlawful content.

The law applies to electronic service providers which, for profit-making purposes, operate Internet platforms which are designed to enable users to share any content with other users or to make such content available to the public (ie social media).[42] The law specifically exempts from its application platforms offering journalistic or editorial content as well as platforms which are designed to enable individual communication or the dissemination of specific content.[43] Furthermore, the NetzDG does not apply to social media with less than 2 million registered users in Germany. Moreover, the NetzDG covers unlawful content under Section 1(3) which refers to criminal law provisions established by the German criminal code,[44] including defamation, dissemination of propaganda, public incitement to crime and hate speech.

Several provisions in this law concern transparency.[45] Social media receiving more than 100 complaints per year must prepare every six months a report, in German, disclosing information about the handling of the complaints procedure.[46] Moreover, providers are also obliged to publish these reports in the Federal Gazette and on their own website no later than one month after the half-year concerned has ended.[47]

---

[41] Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken (NetzDG), 1 September 2017, <https://germanlawarchive.iuscomp.org/?p=1245>.
[42] NetzDG, Art 1.
[43] Ibid.
[44] Ibid. Art 1(3). The NetzDG refers to the following Sections of the criminal code: 86, 86a, 89a, 91, 100a, 111, 126, 129 to 129b, 130, 131, 140, 166, 184b in connection with 184d, 185 to 187, 241 or 269.
[45] Ibid. Art 2.
[46] Ibid. Art 2(2). The law provides a minimum list: '1) General observations outlining the efforts undertaken by the provider of the social network to eliminate criminally punishable activity on the platform; 2) description of the mechanisms for submitting complaints about unlawful content and the criteria applied in deciding whether to delete or block unlawful content; 3) number of incoming complaints about unlawful content in the reporting period, broken down according to whether the complaints were submitted by complaints bodies or by users, and according to the reason for the complaint; 4) organisation, personnel resources, specialist and linguistic expertise in the units responsible for processing complaints, as well as training and support of the persons responsible for processing complaints; 5) membership of industry associations with an indication as to whether these industry associations have a complaints service; 6) number of complaints for which an external body was consulted in preparation for making the decision; 7) number of complaints in the reporting period that resulted in the deletion or blocking of the content at issue, broken down according to whether the complaints were submitted by complaints bodies or by users, according to the reason for the complaint, according to whether the case fell under section 3 subsection (2) number (3) letter (a), and if so, whether the complaint was forwarded to the user, and whether the matter was referred to a recognised self-regulation institution pursuant to section 3 subsection (2) number (3) letter (b); 8) time between complaints being received by the social network and the unlawful content being deleted or blocked, broken down according to whether the complaints were submitted by complaints bodies or by users, according to the reason for the complaint, and into the periods 'within 24 hours'/'within 48 hours'/'within a week'/'at some later point'; 9) measures to inform the person who submitted the complaint, and the user for whom the content at issue was saved, about the decision on the complaint'.
[47] Ibid, Art 2(1). The reports published on their own website shall be easily recognisable, directly accessible an

Besides, the NetzDG requires social media to put in place and maintain an easily accessible, effective and transparent procedure for handling complaints about unlawful content.[48] Social media are required to train those managing complaints and set monthly monitoring checks of procedures involving the handling of complaints by social media management. Among these obligations, it is worth focusing on the obligation for social media to remove and block content.[49] Where content is 'manifestly unlawful', social media are required to remove it within 24 hours of receiving the complaint unless the social network has reached an agreement with the competent law enforcement authority on a longer period for deleting or blocking any manifestly unlawful content.[50] For content that is not manifestly unlawful, the NetzDG provides additional seven days to investigate the content at stake.[51] Even this term can be extended when: a) 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; in such cases, the social network can give the user an opportunity to respond to the complaint before the decision is rendered; b) the social network refers the decision regarding unlawfulness to a recognised self-regulation institution within 7 days of receiving the complaint and agrees to accept the decision of that institution.[52]

Failure to comply with these provisions can lead to the impositions to fines up to 5 million euro, and some offences may be sanctioned even if not committed in the Federal Republic of Germany.[53]

3.2 Partly Free States

Singapore

Singapore's Parliament passed the Protection from Online Falsehoods and Manipulation Act in May 2019 amid harsh criticism from civil society, academia and internet platforms for its far-reaching effects.[54] This legislation targets content that is 'false or misleading, whether wholly or in part' and/or there are reasons to believe it affects public interest.[55] It introduced a graduated approach – ranging from corrections to content takedown – to protect public interest, loosely defined. Although the draft

d permanently available.

[48] Ibid, Art 3(1). According to Art 3(5): 'The procedures in accordance with subsection (1) may be monitored by an agency tasked to do so by the administrative authority named in section 4'.

[49] Just the day after the entry into force of this regulation, Beatrix von Storch, the deputy leader of the Alternative for Germany (AfD) party, was suspended from both Twitter and Facebook for an anti-Muslim message she had posted on New Year's Eve. Philip Oltermann and Pádraig Collins 'Two members of Germany's far-right party investigated by state prosecutor' *The Guardian* (2 January 2018) <https://www.theguardian.com/world/2018/jan/02/german-far-right-mp-investigated-anti-muslim-social-media-posts>.

[50] Ibid, Art 3(2)(2).

[51] Ibid, Art 3(2)(2).

[52] Ibid, Art. 4.

[53] Ibid, Art 4(3).

[54] Protection from Online Falsehoods and Manipulation Bill (2019) <https://sso.agc.gov.sg/Bills-Supp/10-2019/Published/20190401?DocDate=20190401>.

[55] Ibid, Art 2.

legislation was subject to public consultation via the Select Committee on Deliberate Online Falsehoods, the hearings appeared to have only been a formality, without a real impact on the drafting process or the final text.[56]

The stated goal of this law is to 'prevent the electronic communication in Singapore of false statements of fact, to suppress support for and counteract the effects of such communication, to safeguard against the use of online accounts for such communication and for information manipulation, to enable measures to be taken to enhance transparency of online political advertisements, and for related matters'.[57] The prohibition of communication of 'false statements of fact' in Singapore applies to both individuals and online intermediaries in or outside the country for statements likely to be prejudicial to the security, the public health, public safety, public tranquillity or public finances; or to friendly relations of Singapore with other countries or to influence the outcome of an election or a referendum, incite feelings of enmity, hatred or ill-will between different groups of persons; or diminish public confidence in the performance of any duty or function of, or in the exercise of an power by public authorities.[58]

Its broad scope of provisions includes prohibitions for disseminating statements known to be false or having reason to believe so, as well as using or creating inauthentic online accounts or bots to do so. Ministers are empowered by this law to require that a competent authority takes a set of measures – from targeted corrections to access blocking orders issued to an internet access provider to take reasonable steps to disable access by end-users in Singapore to a specified online location. For the individuals, the sanctions consist in fines from S$ 20,000 to S$ 100,000 and/or imprisonment from 1 to 10 years, whereas for intermediaries they generally range between S$ 500,000 to S$ 1 million. The law extends to online platforms, traditional media outlets and broadcasters.

An appeal to a High Court can only be made after the person/intermediary has first applied to the Minister to vary or cancel it and the Minister has refused in whole or in part. Importantly, specific platforms or outlets can become 'declared online locations' once 3 or more active measures in the scope of this law have been communicated in Singapore and at least 3 of those within 6 months before the date the Declaration is made. There is a requirement to the owner or operator to inform end-users about the Declaration, but also a general prohibition on providing financial support to declared online locations, as follows: 'A prescribed digital advertising intermediary or prescribed internet intermediary must take reasonable steps (both in and outside Singapore) to ensure that it does not, when acting as a digital advertising intermediary or an internet intermediary, facilitate the communication in Singapore of any paid content that gives publicity to or otherwise promotes an online location that includes the statement or material subject to Part 3 Direction of Part 4 Direction.[59]

The law allows the local courts to extend decisions and impose sanctions and correction measures beyond the borders of the city-state, to eliminate effects in Singapore.

---

[56] See <https://singaporecan.wordpress.com/2018/04/02/civil-society-activists-criticise-singapores-select-committee-hearings/>.
[57] Protection from Online Falsehoods and Manipulation Act, Preamble.
[58] Ibid, Art 7.
[59] Ibid, Art 47.

Kenya

The Kenyan Computer Misuse and Cybercrimes Act has a wide security-focused scope, from unauthorized access, computer fraud and cyber espionage to child pornography, cyber harassment and false publications.[60] It was enacted by the President of Kenya in May 2018 as an 'act of Parliament to provide for offences relating to computer systems; to enable timely and effective detection, investigation and prosecution of computer and cybercrimes; to facilitate international co-operation in dealing with computer and cybercrime matters; and for connected purposes'.[61] Falsehood and misinformation are punished and the law states: 'A person who intentionally publishes false, misleading or fictitious data or misinforms with intent that the data shall be considered or acted upon as authentic, with or without any financial gain, commits an offence and shall, on conviction, be liable to a fine not exceeding five million shillings or to imprisonment for a term not exceeding two years, or to both'.[62] Relatedly, another provision could also be relevant: 'A person who intentionally inputs, alters, deletes, or suppresses computer data, resulting in inauthentic data with the intent that it be considered or acted upon for legal purposes as if it were authentic, regardless of whether or not the data is directly readable and intelligible commits an offence and is liable, on conviction, to fine not exceeding ten million shillings or to imprisonment for a term not exceeding five years, or to both'.[63]

The criminalization of content-related offences does not come after an assessment of the dishonest intent or harm done. At the same time, in certain situations, it might be difficult to operate with broadly defined categories such as 'false, misleading, fictitious data', as it might include controversial content. In the absence of a universal 'truth' that could be determined by the authorities, there is potential for abusing such provisions to curtail investigative journalism or creative writing. The law marginally covers falsehood and misinformation-related legal remedies and appeal system, focusing extensively on criminal investigations for security, espionage and interference.

3.3 Not Free States

China

China had introduced measures to tackle disinformation before the global rise of the 'fake news' debate.[64] On September 25, 2000, the State Council issued a regulation, the Administrative Measures on Internet Information Services,[65] stating that producing, reproducing, publishing, or spreading

[60] Computer Misuse and Cybercrimes Act (2018) <http://kenyalaw.org/kl/fileadmin/pdfdownloads/Acts/ComputerMisuseandCybercrimesActNo5of2018.pdf>.
[61] Ibid, Preamble.
[62] Ibid, Art 12.
[63] Ibid, Art 14(1).
[64] Maria Repnikova, 'China's Lessons for Fighting Fake News' *Foreign Policy* (6 September 2018) <https://foreignpolicy.com/2018/09/06/chinas-lessons-for-fighting-fake-news/>.
[65] State Council, Administrative Measures on Internet Information Services, 25 September 2000, <http://www.gov.cn/gongbao/content/2000/content_60531.htm (in Chinese), archived at https://perma.cc/M6J4-HV7V>.

prescribed information content, including rumours that disrupt social order or undermines social stability, is a crime.[66] Where service providers discover that this information is transmitted or published on their spaces, they must immediately stop the transmission, keep the relevant records, and report the matter to competent government authorities.[67]

In 2013, China threatened to sanction users with up to seven years in prison for posting unverified information, if it gets viewed 5,000 times or shared more than 500 times.[68] In 2015, China's National People's Congress Standing Committee adopted the Ninth Amendment to the Criminal Law of the People's Republic of China, criminalising the spread of fake news that seriously disturbs public order through an information network or other media and punishing this conduct with up to seven years of imprisonment.[69]

Between 2016 and 2017, China criminalized manufacturing or spreading rumours undermining economic and social order,[70] and adopted a law called Provisions for the Administration of Internet News Information Services requiring online-news providers to reprint information of public officials without distortions or falsehoods and punishing the publication of false information with fines. Furthermore, where a crime is committed, criminal penalties also apply according to the law.[71] In 2018, China adopted another regulation requiring microblogging service providers to establish mechanisms to highlight and tackle rumours.

Russia

In March 2019, Russia adopted two laws to tackle disinformation.[72] The new regulation is based on the amendments to existing legislation recognizing broader powers to public authorities to tackle 'fake news',[73] and establishing administrative liability for the dissemination of information that 'expresses contempt for society, the state and official state symbols' via electronic networks.[74]

The aim of the first law is to curb the distribution of 'unreliable information' defined as 'unreliable socially significant information disseminated under the guise of reliable messages, which creates a threat to life and (/or) the health of citizens or property, the threat of mass disturbance of

[66] Ibid, Art 20.
[67] Ibid, Art 16.
[68] Jonathan Kaiman, 'China cracks down on social media with threat of jail for ″online rumours″' The Guardian 10 September 2013 <https://www.theguardian.com/world/2013/sep/10/china-social-media-jail-rumours>.
[69] Ninth Amendment to the PRC Criminal Law, 1 November 2015 <http://www.npc.gov.cn/npc/xinwen/2015-08/31/content_1945587.htm>.
[70] PRC Cybersecurity Law, 7 November 2016, <http://www.npc.gov.cn/npc/xinwen/2016-11/07/content_2001605.htm>.
[71] Provisions on Administration of Internet News Information Services, 2 May 2017 <http://www.cac.gov.cn/2017-05/02/c_1120902760.htm>.
[72] For an overview, see Oreste Pollicino, 'Fundamental Rights as Bycatch – Russia's Anti-Fake News Legislation' *Verfassungsblog.de* (28 March 2019) <https://verfassungsblog.de/fundamental-rights-as-bycatch-russias-anti-fake-news-legislation/>.
[73] Federal Law of 18.03.2019 No. 31-FZ On Amendments to Article 15-3 of the Federal Law on Information, Information Technologies and on Information Protection.
[74] Federal Law of 18.03.2019 No. 30-FZ On Amending the Federal Law on Information, Information Technologies and Information Protection.

public order and (/or) public safety, or the threat of creating or impairing the proper operation of vital elements of transport or social infrastructure, credit institutions, energy facilities, industry or communications'.[75] The Federal Service for Supervision of Communications, Information Technology and Mass Media (*Roskomnadzor*) is the oversight authority, which, based on complaints about 'unreliable information' lodged by the State Prosecution Service either *ex officio* or following a complaint by a third party, has the power to order providers to delete the content at stake. If providers fail to comply with this order within 24 hours, the *Roskomnadzor* can restrict access to the internet.

The second piece of legislation restricts access concerning 'information expressed in an indecent form that offends human dignity and public morality, or displays obvious disrespect for society, the state, the official state symbols of the Russian Federation, the Constitution of the Russian Federation or the bodies exercising state power in the Russian Federation'.[76] In this case, only the General-Prosecutor can lodge a complaint with the *Roskomnadzor* which, after this step, is required to send a notice in Russian and English to the hosting provider, who alerts the content provider. The content provider is obliged to delete the information within 24 hours of receipt of notification from the hosting provider; in the absence of a removal, it is required to limit access to the information resource 24 hours after receipt of Roskomnadzor notification. If the hosting provider does not comply with the request, the Roskomnadzor can order communication operators to limit access to the information source.[77]

Simultaneously, the Federation Council approved the associated law together with amendments to Russia's Code of Administrative Offences, which stipulates liability in the form of penalties of up to 1.5 million rubles (around $23,000) for the spread of untrue and distorting information. Under the 'fake news' law, repeat offenders will face fines of up to 1.5 million rubles – 20,000€, while insult to authorities can cost up to 4,000€ and 15 days in jail.[78] A number of applications of the law have been already reported.[79]

4. Analogies and Differences in the Fight against Disinformation

In the previous section, we outlined the primary regulatory strategies in six countries, providing insights on the relations with other national laws. Beyond territorial borders, it is both timely and relevant to compare these laws to understand analogies and differences between democratic and authoritarian states in tackling online disinformation. The comparison below is done according to the three dimensions used in the analysis, namely (1) sectorial vs general application of the law; (2)

[75] Tass, 'Putin signs law on blocking fake news' 18 March 2019 <https://tass.com/politics/1049186>.
[76] Tass, 'Putin signs law to fight insults to state symbols' 18 March 2019 <https://tass.com/politics/1049204>.
[77] Ibid.
[78] Emily Tamkin, 'With Putin's signature, 'fake news' bill becomes law' *The Washington Post* 18 March 2019 <https://www.washingtonpost.com/world/2019/03/18/with-putins-signature-fake-news-bill-becomes-law/?utm_term=. 34ecad8cc254>
[79] Maxim Edwards, 'Kremlin's new law against 'online disrespect' proves hard to implement', *Advox* 16 July 2019, <https://advox.globalvoices.org/2019/07/16/kremlins-new-law-against-online-disrespect-proves-hard-to-implement/>.

individual or intermediary targeting; (3) criminal or other sanction for failure to comply with legal provisions.

Firstly, concerning the scope of application, it is possible to observe how the laws analysed in this work apply generally rather than focusing on a specific sector. The only sectorial legislation has been adopted by France since the law primarily concern the electoral period. In all the other cases, laws against disinformation do not restrict their scope to a specific case or time period (e.g. election). In Germany, the NetzDG applies to unlawful content falling under the scope of the German criminal code provisions indicated by the NetzDG. Likewise, in other cases, both partly free and not free countries do not apply to a specific sector affected by disinformation but aims to fight disinformation in all cases, including during elections times (in Singapore). In the case of Kenya, the reference to false information and dissemination is contained in one paragraph of the law with a much wider scope, regulating anything from cyber harassment to cyber espionage.

Secondly, the laws presented here have different addressees, but the overwhelming majority have focused on covering individuals and platforms rather than simply regulating online content moderation. Apart from the German and French laws whose scope of application covers exclusively social media, the other countries extend their obligations to natural persons and, usually, also to online intermediaries (e.g. Russia, Singapore). The most relevant example is China which had already put in place regulation to address the dissemination of rumours by individuals and online intermediaries, even before this new regulatory season against the spread of false news began. More recent regulation confirms this trend towards the criminalisation of natural persons for disseminating false information. This is done either in a non-differentiated way (any spread of information by any means) in the Kenyian example or as a graduated approach in Singapore, where designing a bot that spreads false messages entails a doubling of the fine and years in jail associated with simply sharing a message.

Thirdly, regarding the sanctioning mechanism, it is possible to observe how, except for France, all the laws analysed here provide the possibility to apply criminal sanctions for failure to comply with their obligations, in many cases with minimal scrutiny. Although there are differences in the amount of sanctions applicable to natural persons or online intermediaries, the general trend of these laws is to react against the challenges of disinformation through criminal penalties like monetary sanctions (e.g. Germany), or, in some cases, to imprisonment (e.g. Russia, Singapore, Kenya). The authority empowered to take measures against individuals or information intermediaries varies considerably from country to country. While in Singapore the decision to pursue a suspected act of disinformation rests with a Minister, in France, judges are competent to address which content should be considered false and order its removal.

The Freedom House distinction between free, partly free and not free is a useful guide in situating, contextually, the measures taken against online falsehood. Non-democratic countries tend to have a longer history of regulating falsehood, to which the new laws now add an online dimension. Combining the three criteria of our analysis with the degree of freedom of each country, it is possible to represent the current situation according to the conceptual scheme presented in Figure 1 as follows:

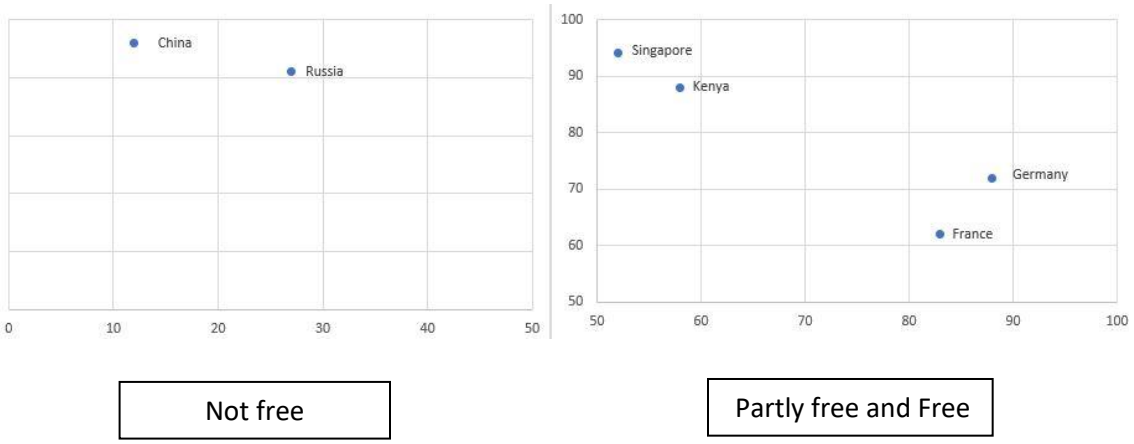

Figure 2. Positioning of the countries analysed according to type of regime and form of regulation

5. Concluding Remarks

This paper analysed new legislation passed to tackle disinformation in six countries around the world, investigating the rationale behind the protections afforded or not to freedom of expression in free, partly free and not free countries. As the spread of fake news made it to the top of the political agenda, both democratic and authoritarian countries felt compelled to respond in order to minimize its effects on advancing the public interest. In many cases, these responses have consisted in new bills and acts of Parliament to sanction the creation, distribution and manipulation of false information.

The primary findings of this study show that the regulatory proposals to counter disinformation around the world are fragmented and often unsatisfactory. Although many of them seem to share the same objective (i.e. fighting disinformation), there is no harmonized approach to tackling the issue. While some countries have adopted limited scope legislation (e.g. France), others have applied restrictions based on potential falsehood for a broad range of activities against public interest. Most of these pieces of legislation use vague definitions of 'public interest', 'public safety', 'falsehood', going as far as encompassing harm done to friendly relationships with other states or to harming the level of trust in public authorities. These divergent regulatory solutions often raise serious concerns for freedom of expression. Without considering the States which have decided not to intervene in the information market (and have relied instead on publishing reports promoting media literacy and forming task forces and working groups to analyse the risk of disinformation), the countries included in this analysis have mostly criminalised the spread of disinformation by sanctioning users. Those that have decided to target both natural persons and companies, usually rely on a mixed set of sanctions including time in prison and substantial fines.

Beyond these trends emerging in regulating disinformation, this research pinpoints that authoritarian states are not the only countries to have adopted restrictive measures to control the spread of messages in the digital environment. Indeed, some democratic states have approached the issue of disinformation by mirroring an authoritarian approach, imposing limitations on freedom of expression, with long-term chilling effects. Nevertheless, it is worth underlining how established democracies have proved to be very cautious in taking steps to regulate disinformation, arguably for concerns over the preservation of democratic values and civil liberties. Disinformation

countermeasures with limited legal remedies are especially widespread in Africa and in Asia, where they tend to combine with increased surveillance of online activities.

These findings represent a non-exhaustive preview of the state-driven actions to fight disinformation and provide new insights into the dangers of fast-paced regulation on highly politicized topics. Further research needs to focus on the application of these laws and their effects in the short- and medium-term. Moreover, the relationship with the private sector, in particular the delegation of responsibility for restricting access to content online, needs to be further interrogated. As oversight mechanisms start to be introduced, it is crucial to understand to what extend they serve the state and to what extent they serve the larger public. In this paper, we show that the disinformation arena is a test case for the resilience of democratic systems in times of alleged decay. The similarities we note between authoritarian and democratic regimes provide a cautionary tale.