

## Fake News in Times of Pandemic and Beyond: An Enquiry into the Rationales for Regulating Information Platforms

MIRA BURRI\*

*The COVID-19 pandemic threw our societies in dire times with deep effects on all societal sectors and on our lives. The pandemic was accompanied by another phenomenon also associated with grave consequences – that of the ‘infodemic’. Fake news about the cause, prevention, impact and potential cures for the coronavirus spread on social platforms and other media outlets, and continue to do so. The chapter takes this infodemic as a starting point to exploring the broader phenomenon of online misinformation. The legal analysis in this context focuses on the rationales for regulating Internet platforms as critical information intermediaries in a global networked media space. As Internet platforms do not fall under the category of media companies, they are currently not regulated in most countries. Yet, the pressure to regulate them, also in light of other negative phenomena, such as hate speech proliferation, political disinformation and targeting, has grown in recent years. The regulatory approaches differ, however, across jurisdictions and encompass measures that range from mere self-regulatory codes to more binding interventions. Starting with some insights into the existing technological means for mediating speech online, the power of platforms, and more specifically of their influence on the conditions of freedom of expression, the chapter discusses, in particular, the regulatory initiatives with regard to information platforms in the United States and in the European Union, as embedded in different traditions of free speech protection. The chapter offers an appraisal of the divergent US and EU approaches and contemplates the adequate design of regulatory intervention in the area of online speech in times of infodemic and beyond it.*

### A. Introduction: The Changing Landscape of Fake News

The COVID-19 pandemic has had disastrous effects on human health and well-being as well as on the economy and other areas of societal life. One factor that played into this global crisis has been the lack of trustworthy and reliable information sources and the simultaneous inflow of misinformation that fed into unfortunate individual choices, shared conspiracy theories<sup>1</sup> (such as myths about 5G installations spreading COVID-19 or a particular ethnic or religious group being at the virus’ origin) and patterns of group resistance.<sup>2</sup> The novelty of the COVID virus

---

\* Professor of International Economic and Internet Law, University of Lucerne. Contact: [mira.burri@unilu.ch](mailto:mira.burri@unilu.ch). All websites have been last accessed on 25 August 2021.

<sup>1</sup> On the origins and the reason for distribution of conspiracy theories, see e.g. Cass R. Sunstein and Adrian Vermeule, ‘Conspiracy Theories: Causes and Cures’, *The Journal of Political Philosophy* 17 (2009), 202–227.

<sup>2</sup> For an overview of the different misinformation threats, see e.g. European Commission and the High Representative of the Union for Foreign Affairs and Security Policy, *Tackling COVID-19 Disinformation: Getting the Facts Right*, JOIN(2020) 8 final, 10 June 2020; see also J. Scott Brennen et al., *Types, Sources, and Claims of COVID-19 Misinformation* (Oxford: Reuters Institute for the Study of Journalism, 2020); Darrin Baines and Robert J. R. Elliott, ‘Defining Misinformation, Disinformation and Malinformation: An Urgent Need for Clarity during the COVID-19 Infodemic’, *University of Birmingham Department of Economics Discussion Papers* 20-06

and the related gaps in knowledge created an almost ideal breeding ground for false or misleading narratives to spread.<sup>3</sup> Pursuant to the World Health Organization (WHO), such an infodemic creates confusion and distrust, and ultimately undermines an effective public health response.<sup>4</sup> There are numerous studies that show how exponentially this infodemic has spread on different social platforms and how real its implications have been.<sup>5</sup>

Admittedly, fake news, including ones about the origins and spread of diseases, are not new: '[m]isinformation, disinformation and propaganda have been features of human communication since at least the Roman times',<sup>6</sup> and over the ages, they have only been facilitated by technological advances starting with the invention of the printing press and now being immensely expedited by the Internet as a global communication platform.<sup>7</sup> While there are different narratives and timelines on the beginnings and the evolution of fake news as a societal phenomenon,<sup>8</sup> for the purpose of this chapter's discussion, it is critical to highlight what is specific about the nature of contemporary fake news: 'There are three unique parts to modern fake news that make it different from older varieties of intentionally exaggerated or false reporting: the who, the what, and the how'.<sup>9</sup> With regard to the 'who', what is peculiar is that today's fake news are not produced by classic media outlets, although they may be taken up by them.<sup>10</sup> The origins of misinformation rather lead back to governments driven by ideological interests or tech-savvy individuals looking for a certain reward. With regard to the 'what' of

---

(2020); Adam M. Enders, 'The Different Forms of COVID-19 Misinformation and Their Consequences', *The Harvard Kennedy School (HKS) Misinformation Review* (2020).

<sup>3</sup> See European Commission, *supra* note 2.

<sup>4</sup> WHO stated that 'infodemics are an excessive amount of information about a problem, which makes it difficult to identify a solution. They can spread misinformation, disinformation and rumours during a health emergency. Infodemics can hamper an effective public health response and create confusion and distrust among people'. See WHO, Coronavirus disease 2019 (COVID-19), Situation Report 45, 5 March 2020, available at: [https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200305-sitrep-45-covid-19.pdf?sfvrsn=ed2ba78b\\_4](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200305-sitrep-45-covid-19.pdf?sfvrsn=ed2ba78b_4); see also Managing the COVID-19 Infodemic: Promoting Healthy Behaviours and Mitigating the Harm from Misinformation and Disinformation, Joint Statement by WHO, UN, UNICEF, UNDP, UNESCO, UNAIDS, ITU, UN Global Pulse, and IFRC, 23 September 2020, available at: <https://www.who.int/news/item/23-09-2020-managing-the-covid-19-infodemic-promoting-healthy-behaviours-and-mitigating-the-harm-from-misinformation-and-disinformation>.

<sup>5</sup> See e.g. Aleksi Knuutila et al., Covid-Related Misinformation on YouTube: The Spread of Misinformation Videos on Social Media and the Effectiveness of Platform Policies, Oxford Internet Institute, COMPROM Data Memo 2020.6, 21 September 2020; Brennen et al., *supra* note 2; Center for Countering Digital Hate, *Malgorithm How Instagram's Algorithm Publishes Misinformation and Hate to Millions during a Pandemic* (London and Washington, DC: Center for Countering Digital Hate, 2020); Rasmus Kleis Nielsen et al., *Communications in the Coronavirus Crisis: Lessons for the Second Wave* (Oxford: Reuters Institute for the Study of Journalism, 2020).

<sup>6</sup> Julie Posetti and Alice Matthews, A Short Guide to the History of 'Fake News' and Disinformation (Washington, DC: International Center for Journalists, 2020), at 1. For understanding the different types of fake news, making a difference between disinformation, misinformation and malinformation, see Claire Wardle, *Understanding Information Disorder* (London: First Draft, 2019); for a slightly different classification, see Edson C. Tandoc Jr., Zheng Wei Lim and Richard Ling, 'Defining "Fake News"', *Digital Journalism* 6 (2018), 137–153.

<sup>7</sup> See e.g. Cherilyn Ireton and Julie Posetti (eds), *Journalism, 'Fake News' and Disinformation* (Paris: UNESCO, 2018) and the next section.

<sup>8</sup> See e.g. Posetti and Matthews, *ibid.*; Jacob Soll, 'The Long and Brutal History of Fake News', *POLITICO Magazine*, 18 December 2016, available at: <http://politi.co/2FaV5W9>; Mike Wendling, 'The (Almost) Complete History of "Fake News"', BBC News, 22 January 2018, available at: <https://www.bbc.co.uk/news/blogs-trending-42724320>; Merriam-Webster, 'The Real Story of "Fake News"', available at: <https://www.merriam-webster.com/words-at-play/the-real-story-of-fake-news>; Steven Poole, 'Before Trump: The real History of Fake News', *The Guardian*, 22 November 2019, available at: <https://www.theguardian.com/books/2019/nov/22/factitious-taradiddle-dictionary-real-history-fake-news>.

<sup>9</sup> Center for Information Technology and Society – UC Santa Barbara, *A Brief History of Fake News* (Santa Barbara, CA: Center for Information Technology and Society, 2019).

<sup>10</sup> *Ibid.*

fake news, it is not only the content that may be distorted but also the source. Deep fake technology has only facilitated the production of pieces of content that appear to be entirely truthful and authentic.<sup>11</sup> Finally, the ‘how’ is quite different, as online platforms are key in the distribution of fake news and their seamless embeddedness in communication processes. Social media tend to be ‘source-agnostic’, as they collect and present news stories from a variety of outlets, regardless of the quality and reliability of the original source. Also, because news stories often reach the users through a certain circle of friends or people they follow, there is a sort of endorsement of the story. Finally, and critically, through the inherent functionality of social platforms’ algorithms, the popularity of fake news can be significantly enhanced and their outreach magnified.<sup>12</sup>

The infodemic and the fake news phenomenon, in general, have become well acknowledged by policy-makers and legislators around the world and have spurred a number of initiatives that try to address them,<sup>13</sup> which have been linked to larger projects of regulating platforms that relate to other negative phenomena, such as hate speech proliferation or extreme opinion building. The following section explores these initiatives and the surrounding debates by first providing a look into the changing role of platforms in the new digital media space and the associated dangers for individual free speech, as well as for a healthy public discourse as a fundament of a democratic society.

## B. Regulating Online Platforms

### I. Understanding Platforms as Information Intermediaries

The Internet has enabled instantaneous sharing of information and communication among millions of people, with a relatively low threshold for participation and seemingly no barriers.<sup>14</sup> Many have hoped that this digital space would, even without the need for state intervention, create the conditions necessary for individual freedom of speech, both in its active and passive dimensions, for a sustainable public discourse and content diversity to flourish.<sup>15</sup> Sadly, this brave new world of the ‘marketplace of ideas’<sup>16</sup> has not materialized so far; instead, the digital space has brought discrete new challenges with it, some of which may call for deliberate action.

A distinct new phenomenon that has captured the attention of both scholars<sup>17</sup> and policy-makers<sup>18</sup> in this debate is the critical role played by platforms.<sup>19</sup> Platforms like social

<sup>11</sup> Ibid. See also Bobby Chesney and Danielle Citron, ‘Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security’, *California Law Review* 107 (2019), 1753–1820.

<sup>12</sup> Center for Information Technology and Society, *supra* note 9.

<sup>13</sup> For updated information on the initiatives around the world, see e.g. <https://infogram.com/covid-19-fake-news-regulation-passed-during-covid-19-1h8n6md5q9kj6xo>.

<sup>14</sup> See e.g. Yochai Benkler, *The Wealth of Networks: How Social Production Transforms Markets and Freedom* (New Haven, CT: Yale University Press, 2006); Richard S. Whitt, ‘A Deference to Protocol: Fashioning a Three-dimensional Public Policy Framework for the Internet Age’, *Cardozo Arts and Entertainment Law Journal* 31 (2013), 689–768.

<sup>15</sup> See e.g. Benkler, *ibid.*; Lawrence Lessig, *Remix: Making Art and Commerce Thrive in the Hybrid Economy* (New York: Penguin, 2009).

<sup>16</sup> For a clarification of the term, see below.

<sup>17</sup> See e.g. Jack M. Balkin, ‘Free Speech Is a Triangle’, *Columbia Law Review* 118 (2018), 2011–2055; Julie E. Cohen, ‘Law for the Platform Economy’, *University of California Davis Law Review* 51 (2018), 133–204; Orly Lobel, ‘The Law of the Platform’, *Minnesota Law Review* 101 (2016), 87–166.

<sup>18</sup> See e.g. European Commission, Online Platforms and the Digital Single Market, COM(2016) 288 final, 25 May 2016; UK House of Commons, Disinformation and ‘Fake News’, Final Report of the Digital, Culture, Media and Sport Committee, 2019.

<sup>19</sup> The focus here is placed not upon the physical intermediaries, such as network operators or Internet service

networking sites, search engines and other types of aggregators, often driven by algorithms, have turned into gatekeepers that command substantial powers in contemporary media environments. They have become content curators and profoundly changed the processes of cultural production, distribution, the conditions for access to content, its use and re-use – in many aspects effectively replacing the role of traditional media as a ‘general interest intermediary’<sup>20</sup> but without any public or other regulatory mandate.<sup>21</sup> The implications of this transformation are multiple, many of them are critical for the design of the law. Some domains, such as copyright law, have been substantially adjusted to this new environment in the last couple of decades, whereby intermediaries have become key enforcers of copyright law and subject to specific types of liability.<sup>22</sup> Other areas, such as antitrust law,<sup>23</sup> have been much slower in this adaptation. The same is true for media law, which has traditionally been entrusted with key functions to ensure conditions for opinion formation, public debate, political and cultural engagement, and social cohesion,<sup>24</sup> as discussed in more detail below.

One key implication of the new digital media space that complicates conventional conditions of free speech, as well as the initiatives to regulate it, is the shift from the standard dualist model of state versus speakers-publishers towards a pluralistic model with complex relations along a tringle of actors, where both the audience and the state become highly dependent on platforms.<sup>25</sup> It has been argued that in such a configuration, nation states tend to create different liability regimes for digital companies that may trigger collateral censorship and prior restraint.<sup>26</sup> Social media companies, in their own right, create sophisticated systems of private governance that govern end users arbitrarily and without due process and transparency. At the same time, users are vulnerable to digital surveillance and manipulation, and the intensified datafication of the digital economy only exacerbates this vulnerability.<sup>27</sup> A second fundamental shift in the new

---

suppliers (although these matter too: see e.g. Benkler, *supra* note 14; Laura DeNardis, *Protocol Politics: The Globalization of Internet Governance* (Cambridge: MIT Press, 2009); Brett M. Frishmann, *Infrastructure: The Social Value of Shared Resources* (Oxford: Oxford University Press, 2012) but upon those gatekeepers existing at the applications and the content levels – the so-called ‘choice intermediaries’ or ‘new editors’. See Natali Helberger, ‘Diversity by Design’, *Journal of Information Policy* 1 (2011), 441–469 and Persephone Miel and Robert Farris, *News and Information as Digital Media Come of Age* (Cambridge, MA: The Berkman Center for Internet and Society, 2008).

<sup>20</sup> Cass R. Sunstein, *Republic.com 2.0* (Princeton: Princeton University Press, 2007).

<sup>21</sup> Philip M. Napoli, ‘Persistent and Emergent Diversity Policy Concerns in an Evolving Media Environment’, in Sean A. Pager and Adam Candeub (eds), *Transnational Culture in the Internet Age* (Cheltenham: Edward Elgar, 2012), 165–181; Philip M. Napoli, ‘Social Media and the Public Interest: Governance of News Platforms in the Realm of Individual and Algorithmic Gatekeepers’, *Telecommunications Policy* 39 (2015), 751–760.

<sup>22</sup> See e.g. Matthew Sag, ‘Internet Safe Harbors and the Transformation of Copyright Law’, *Notre Dame Law Review* 93 (2018), 499–564; Gerald Spindler, ‘Copyright Law and Internet Intermediaries Liability’, in Tatiana-Eleni Synodinou et al. (eds), *EU Internet Law in the Digital Era* (Berlin: Springer, 2020), 3–25; Giancarlo Frosio (ed), *Oxford Handbook of Online Intermediary Liability* (Oxford: Oxford University Press, 2020); Mira Burri and Zaira Zihlmann, ‘Intermediaries’ Liability in Light of the Recent EU Copyright Reform’, *Indian Journal of Intellectual Property Law* 11 (2021).

<sup>23</sup> See e.g. Ariel Ezrachi and Maurice E. Stucke, *Virtual Competition: The Promise and Perils of the Algorithm-Driven Economy* (Cambridge, MA: Harvard University Press, 2016); for a brief overview of the issues, see Mira Burri, ‘Understanding the Implications of Big Data and Big Data Analytics for Competition Law: An Attempt for a Primer’, in Klaus Mathis and Avshalom Tor (eds), *New Developments in Competition Behavioural Law and Economics* (Berlin: Springer, 2019), 241–263.

<sup>24</sup> See e.g. Mira Burri, *Public Service Broadcasting 3.0: Legal Design for the Digital Present* (London: Routledge, 2015).

<sup>25</sup> Balkin (2008), *supra* note 17; also Jack M. Balkin, ‘Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation’, *University of California Davis Law Review* 51 (2018a), 1149–1210.

<sup>26</sup> Balkin (2008), *supra* note 17.

<sup>27</sup> Balkin (2008), *ibid.*; Balkin, *supra* note 25; Kate Klonick, ‘The New Governors: The People, Rules, and

media environment has been epitomized by the enhanced editorial functions of digital platforms as ‘choice intermediaries’ that control the choices for content, communication and engagement we make and the possibility for choices we see.<sup>28</sup> A number of issues can be noted in this context: the first one relates to the critical role of technology, of algorithms, as automated filters,<sup>29</sup> aggregators and even content producers.<sup>30</sup> Another important feature that explains the motivation in selecting a particular algorithm design or making a governance decision as to the availability and prominence of content<sup>31</sup> relates to the nature of multi-sided markets inherent to most digital platforms, where they give users free access to certain services on one side of the platform, while also selling the information collected to advertisers and other companies.<sup>32</sup> This model incentivizes the battle for the attention of the many but also for tailored offerings; it makes data collection and data use more intrusive, which bears upon competition and data protection<sup>33</sup> but is yet to be seriously considered in the domain of media law – despite the fact that it deeply impacts information flows and the conditions of free speech.<sup>34</sup>

Thinking about the societal functions of the media in the context of this chapter’s discussion, it can also be that this complex platform-mediated environment engenders certain risks for the animated public sphere and for a vibrant and diverse culture. First, the possible interferences with users’ individual autonomy and freedom of choice need to be acknowledged. As Latzer et al. argue, while filtering reduces search and information costs and facilitates social orientation, it can be ‘compromised by the production of social risks, amongst others, threats to basic rights and liberties as well as impacts on the mediation of realities and people’s future development’.<sup>35</sup> The second worry relates to the impact of tailored consumption on the engagement of the user in political, social and cultural debates. The personalization of the media diet, as based on a distinct profile or previous experience, ‘promotes content that is geographically close as well as socially and conceptually familiar’;<sup>36</sup> it reflects each individual’s interests and biases and provides no information to disrupt preconceptions or prejudices.<sup>37</sup> Hoffman et al. further argue that social media only exacerbate this effect by combining two dimensions of ‘homophily’ – similarity of peers and of content.<sup>38</sup> While these situations have been labelled differently – as

---

Processes Governing Online Speech’, *Harvard Law Review* 131 (2018), 1598–1670.

<sup>28</sup> See Helberger (2011); Miel and Farris (2008), both supra note 19.

<sup>29</sup> For a taxonomy of the different algorithmic filters, see Engin Bozdag, ‘Bias in Algorithmic Filtering and Personalization’, *Ethics and Information Technology* 15 (2013), 209–227.

<sup>30</sup> Philip M. Napoli, ‘On Automation in Media Industries: Integrating Algorithmic Media Production into Media Industries Scholarship’, *Media Industries Journal* 1 (2014), 33–38; also Florian Saurwein et al., ‘Governance of Algorithms: Options and Limitations’, *info* 17 (2015), 35–49; Natali Helberger, Katharina Kleinen-von Königslöw and Rob van der Noll, ‘Regulating the New Information Intermediaries as Gatekeepers of Information Diversity’, *info* 6 (2015), 50–71; Mira Burri, ‘Discoverability of Local, National and Regional Content Online’, A Thought Leadership Paper written for the Canadian Commission for UNESCO and Canadian Heritage, 7 February 2019a.

<sup>31</sup> Klonick, supra note 27.

<sup>32</sup> See e.g. Kenneth A. Bamberger and Orly Lobel, ‘Platform Market Power’, *Berkeley Technology Law Journal* 32 (2018) 32, 1052–1092; Burri (2019), supra note 23.

<sup>33</sup> Burri (2019), *ibid.*; European Data Protection Supervisor (EDPS), Privacy and Competitiveness in the Age of Big Data, 2014; Oreste Pollicino, *Judicial Protection of Fundamental Rights on the Internet* (Oxford: Hart, 2021).

<sup>34</sup> Cohen, supra note 17; Daniel Kreiss and Shannon C. McGregor, ‘Technology Firms Shape Political Communication: The Work of Microsoft, Facebook, Twitter, and Google with Campaigns During the 2016 US Presidential Cycle’, *Political Communication* (2017), 1–23; Yochai Benkler et al., *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics* (Oxford: Oxford University Press, 2018).

<sup>35</sup> Michael Latzer et al. ‘The Economics of Algorithmic Selection on the Internet’, *Media Change and Innovation Working Paper* (2014), at 29–30;

<sup>36</sup> Christian P. Hoffman et al. ‘Diversity by Choice: Applying a Social Cognitive Perspective to the Role of Public Service Media in the Digital Age’, *International Journal of Communication* 9 (2015), 1360–1381, at 1365.

<sup>37</sup> Hoffman et al., *ibid.*; Sunstein (2007), supra note 20.

<sup>38</sup> Hoffman et al., *ibid.*, at 1365.

‘cyber-ghettos’,<sup>39</sup> ‘filter bubbles’<sup>40</sup> or ‘echo chambers’<sup>41</sup> – they all point to a fragmentation of the public discourse, possible polarization of views and augmentation of the impact of unwanted content, such as fake news.<sup>42</sup> From the perspective of media law and policy, one should also acknowledge that in most media policy toolkits, the underlying causal link between diversity in supply and diversity in consumption may be destroyed.<sup>43</sup> Local and national identities and debates, as well as cultural diversity, may be endangered, as local, regional and national content and quality journalism, especially in the domains of news and current affairs, are marginalized online and rendered hard to ‘discover’.<sup>44</sup>

A final element that complicates the conditions of free speech in the era of platforms is their staggering power – vis-à-vis the states (both domestic and foreign regulators), vis-à-vis other companies on the same or adjacent markets and ultimately vis-à-vis the users. Indeed, it has been argued that platforms have become the ‘new governors’<sup>45</sup> or the ‘emergent transnational sovereigns’ of the digital space.<sup>46</sup> Yet, and as mentioned earlier, this power is often unchecked, and platforms moderate speech practice and cultural communication and engagement with any accountability - neither to their users nor to state agencies.<sup>47</sup> Against this backdrop of the new digital space and its increasing platformization, the next sections explore the initiatives that have been developed in the past few years to regulate platforms in diverse ways, with a specific focus laid upon the diverging approaches of the United States (US) and the European Union (EU) in this context.

## II. Regulatory Initiatives Addressing the Platformization of the Media Space and Online Misinformation

### 1. Introduction

The developments around platformization of the communication environment and the risks brought with it do not occur in a regulatory vacuum. They are embedded in long-standing traditions of international human rights protection,<sup>48</sup> as well as in domestic constitutional traditions and sophisticated frameworks of secondary legislation and judicial precedent.<sup>49</sup> This

<sup>39</sup> Peter Dahlgren, ‘The Internet, Public Spheres, and Political Communication’, *Political Communication* 22 (2005), 147–162.

<sup>40</sup> Eli Pariser, *The Filter Bubble: What the Internet Is Hiding from You* (New York: Viking, 2011).

<sup>41</sup> Cass R. Sunstein, *Echo Chambers: Bush v. Gore Impeachment, and Beyond* (Princeton: Princeton University Press, 2001); Sunstein (2007), supra note 20; Cass R. Sunstein, *Going to Extremes: How Like Minds Unite and Divide* (Oxford: Oxford University Press, 2009).

<sup>42</sup> Filter bubbles, together with ‘information cascades’ and the human attraction to negative and novel information have been said to fuel the distribution and virality of fake news. For a careful analysis of these phenomena of online communication, see Chesney and Citron, supra note 11, in particular at 1765–1768.

<sup>43</sup> Ellen P. Goodman, ‘Media Policy Out of the Box: Content Abundance, Attention Scarcity, and the Failures of Digital Markets’, *Berkeley Technology Law Journal* 19 (2004), 1389–1472; Natali Helberger, ‘Exposure Diversity as a Policy Goal’, *Journal of Media Law* 4 (2012), 65–92; Napoli (2012), supra note 21; Philip M. Napoli et al., *Assessing Local Journalism: News Deserts, Journalism Divides, and the Determinants of the Robustness of Local News* (New Brunswick, NJ: News Measures Research Project, 2018).

<sup>44</sup> Napoli et al. (2018), *ibid.*; Burri (2019a), supra note 30; Fenwick McKelvey and Robert Hunt, ‘Discoverability: Toward a Definition of Content Discovery through Platforms’, *Social Media + Society*, January/March (2019), 1–15.

<sup>45</sup> Balkin (2018), supra note 17; Klonick (2018), supra note 27.

<sup>46</sup> Cohen, supra note 17.

<sup>47</sup> Balkin (2018a), supra note 25; see also David Kaye, *Speech Policy: The Struggle to Govern the Internet* (New York: Columbia Global Reports, 2019).

<sup>48</sup> See e.g. Olivier De Schutter, *International Human Rights Law*, 2nd edn. (Cambridge: Cambridge University Press, 2014).

<sup>49</sup> See e.g. Daniel A. Farber, *The First Amendment: Concepts and Insights* (St. Paul, MN: Foundation Press, 2019);

has been clearly recognized with regard to fake news, as several key international and regional human rights bodies, confirmed the applicability of existing frameworks in that: ‘States may only impose restrictions on the right to freedom of expression in accordance with the test for such restrictions under international law, namely that they be provided for by law, serve one of the legitimate interests recognised under international law, and be necessary and proportionate to protect that interest’<sup>50</sup> and that ‘[g]eneral prohibitions on the dissemination of information based on vague and ambiguous ideas, including “false news” or “non-objective information”, are incompatible with international standards for restrictions on freedom of expression [...] and should be abolished’.<sup>51</sup>

Yet, it should also be noted that free speech law has not been harmonized sufficiently at the international level, and we have, at times, profoundly different regimes in different countries. As platforms act as global players across jurisdictions and considering the borderless nature of the Internet,<sup>52</sup> the legal design that should appropriately safeguard fundamental rights and values within the sovereign state is only made more difficult.<sup>53</sup> In addition, most, if not all, of the platforms are American and as such tend to govern, both by technological means and human intervention, reflecting the First Amendment principle, social and corporate responsibility and the liability exceptions under US law.<sup>54</sup> In the following sections, we briefly trace the differences in the US and EU standards of free speech protection and subsequently expound on the initiatives undertaken to regulate free speech on platforms, and fake news in particular, in these major jurisdictions. While the initiatives to fight misinformation can be put into different categories covering a great variety of responses, such as (i) identification responses (including monitoring, fact-checking and investigative responses); (ii) ecosystem responses aimed at producers and distributors (including legislative and policy responses, counter-disinformation campaigns and electoral-specific responses); (iii) responses within production and distribution (including curatorial, technical, demonetization and advertising-linked responses); and (iv) responses aimed at target audiences (including normative ethical and educational responses;

---

Noah R. Feldman and Kathleen M. Sullivan, *First Amendment Law*, 7th edn. (St. Paul, MN: West Academic, 2019); Paul Keller, *European and International Media Law: Liberal Democracy, Trade, and the New Media* (Oxford: Oxford University Press, 2011); Jan Oster, *European and International Media Law* (Cambridge: Cambridge University Press, 2017); Pollicino, supra note 33.

<sup>50</sup> 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 Freedom of Expression and ‘Fake News’, Disinformation and Propaganda*, FOM.GAL/3/17, 3 March 2017, at 1. *General Principles*, para. (a).

<sup>51</sup> *Ibid.*, at 2. *Standards on Disinformation and Propaganda*, para. (a).

<sup>52</sup> See e.g. Kristin E. Eichensehr, ‘The Cyber-Law of Nations’, *The Georgetown Law Journal* 103 (2015), 317–380; Jennifer Daskal, ‘The Un-Territoriality of Data’, *The Yale Law Journal* 125 (2016); 326–398; Kristin E. Eichensehr, ‘Data Extraterritoriality’, *Texas Law Review* 95 (2017), 145–160.

<sup>53</sup> See e.g. the *Yahoo!* case as one of the first on free speech violation and online jurisdiction. There the US court when faced with the recognition and implementation of the French court order under the ‘comity of nations’ doctrine stated that: ‘Absent a body of law that establishes international standards with respect to speech on the Internet and an appropriate treaty or legislation addressing enforcement of such standards to speech originating within the United States, the principle of comity is outweighed by the Court’s obligation to uphold the First Amendment’ (see *Yahoo! II*, 169 F. Supp. 2d, at 1193). For more on the case, see Jack Goldsmith and Tim Wu, *Who Controls the Internet? Illusions of a Borderless World* (Oxford: Oxford University Press, 2001), at 49–64; Marc H. Greenberg, ‘A Return to Lilliput: The *LICRA v. Yahoo* – Case and the Regulation of Online Content in the World Market’, *Berkeley Technology Law Review* 18 (2003), 1191–1258.

<sup>54</sup> See Klonick (2018), supra note 27; see also Anupam Chander, ‘How Law Made Silicon Valley’, *Emory Law Journal* 63 (2014), 639–694.

empowerment and credibility labelling responses),<sup>55</sup> the focus of this chapter is solely on the legislative responses with regard to platforms. It should also be noted that the chapter's enquiry is by no means comprehensive but is merely illustrative, as there is a flurry of initiatives around the world with currently around 28 countries having passed legislation related to disinformation, either through new laws or by updating existing regulations in the areas of media, electoral, cybersecurity or criminal law.<sup>56</sup>

## 2. *Developments in the United States*

To begin with, it is fair to note that there have not been any major regulatory projects in the United States that seek to address online misinformation. And the developments that we have seen in the past years with regard to online platforms have been in the field of self-regulation, with major companies like Facebook or Twitter designing a palette of measures for their platforms to tackle misinformation in particular and bad speech in general.<sup>57</sup> This is not surprising, as the protection of free speech in the United States is in many ways specific when compared with other countries and awarded a higher value.<sup>58</sup> For starters, the First Amendment to the US Constitution clearly states that 'Congress shall make no law [...] abridging the freedom of speech'. Whereas there have been some restrictions on free speech, as later discussed, these are fairly limited in scope, and the doctrine of 'the marketplace of ideas' as endorsed by Justice Oliver Wendell Holmes,<sup>59</sup> which sees free speech as an open marketplace where ideas compete against each another for acceptance by the public, has been sustained. The Internet has not been seen to change the nature of the First Amendment's application, as confirmed by the very first Internet-related case by the US Supreme Court in *Reno v. ACLU*,<sup>60</sup> where the Court pointed out that 'the content on the Internet is as diverse as human thought' and established 'no basis for qualifying the level of First Amendment scrutiny that should be applied to this medium'.<sup>61</sup>

For the legal test in defining the scope of the First Amendment, as developed over the years, courts need to determine (i) whether the law regulates a category of speech that is unprotected under the First Amendment or granted lesser than full protection, which gives the government certain regulatory authority, and (ii) whether the law is a content-based restriction, which is then presumed invalid under strict scrutiny, or a content-neutral restriction, which is subject to

---

<sup>55</sup> See Kalina Bontcheva and Julie Posetti (eds), *Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression*, Broadband Commission research report on 'Freedom of Expression and Addressing Disinformation on the Internet' (Geneva/Paris: ITU/UNESCO, 2020), in particular at 36–40.

<sup>56</sup> Bontcheva and Posetti, *ibid.*, at 108; see also Judit Bayer et al., *Disinformation and Propaganda – Impact on the Functioning of the Rule of Law in the EU and its Member States*, Study for the European Parliament (Brussels: European Parliament, 2019); Peter Roudik et al., *Initiatives to Counter Fake News in Selected Countries: Argentina, Brazil, Canada, China, Egypt, France, Germany, Israel, Japan, Kenya, Malaysia, Nicaragua, Russia, Sweden, United Kingdom* (Washington, DC: The Law Library of Congress, 2019).

<sup>57</sup> See e.g. Dawn C. Nunziato, 'The Marketplace of Ideas Online', *Notre Dame Law Review* 94 (2019), 1519–1583; Kate Klonick, 'The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression', *The Yale Law Journal* 129 (2020), 2418–2299.

<sup>58</sup> See e.g. Claudia E. Haupt, 'Regulating Hate Speech: Damned If You Do and Damned If You Don't – Lessons Learned from Comparing the German and U.S. Approaches', *Boston University International Law Journal* 23 (2005), 300–335 (also providing an overview of the comparative literature); Ioanna Tourkochoriti, 'Speech, Privacy and Dignity in France and in the U.S.A.: A Comparative analysis', *Loyola L.A. International and Comparative Law Review* 38 (2016), 101–182.

<sup>59</sup> *Abrams v. United States*, 250 U.S. 616 (1919) (dissenting opinion Holmes).

<sup>60</sup> 521 U.S. 844 (1996).

<sup>61</sup> *Ibid.*, at para. 885. In the more recent case of *Packingham v. North Carolina*, 137 S. Ct. 1730 (2017), the Supreme Court compared social media platforms to a town square and recognized their function as a forum to exchange ideas and viewpoints.

intermediate scrutiny, a less speech-protective test.<sup>62</sup> The content-based restriction never passes the test, as ‘... the government has no power to restrict expression because of its message, its ideas, its subject matter, or its content’<sup>63</sup> and the state may not prohibit the expression of an idea ‘simply because society finds the idea itself offensive or disagreeable’.<sup>64</sup> In the application of strict scrutiny, a law can ‘survive’ only when it is narrowly tailored to promote a compelling government interest.<sup>65</sup> Restrictions on speech that have been upheld so far relate to a limited number of situations, such as obscenity, child pornography, fraud, treason, incitement to crime, fighting words, true threats, and speech presenting some grave and imminent threat the government has the power to prevent,<sup>66</sup> whereby there must be a realistic, factual assessment of harm.<sup>67</sup> The test with regard to defamation and libel is also hard to pass since *New York Times Co. v. Sullivan*, where the Court found that a public official seeking to recover damages for a defamatory falsehood relating to his official conduct must prove that the statement was made with ‘actual malice’ – that is, with the knowledge that it was false or with reckless disregard of whether it was false or not.<sup>68</sup>

Since the 2012 Supreme Court decision in *United States v. Alvarez*,<sup>69</sup> it has also become clear that false statements fall within the scope of the First Amendment. In *Alvarez*, the Court found the 2005 Stolen Valor Act, which criminalized a falsely claimed receipt of military decorations or medals, unconstitutional, and the majority highlighted that punishing false speech would deter an open and vigorous expression of views<sup>70</sup> and that less restrictive measures, such as counter-speech, could promote the state’s legitimate interests.<sup>71</sup> This said, some US states do have false reporting statutes for very specific situations imposing criminal liability for false speech related to emergencies or natural catastrophes, with New York’s false reporting statute<sup>72</sup> being the most far-reaching in this regard, as it does not require knowledge or intent with respect to the ensuing public alarm or inconvenience.<sup>73</sup> Next to this robust protection of free speech granted in the United States, which very often is given primacy over other rights, such as privacy,<sup>74</sup> the US Speech Act bans enforcement of judgements that would violate the free

<sup>62</sup> Louis W. Tompros, Richard A. Crudo, Alexis Pfeiffer and Rahel Boghossian, ‘The Constitutionality of Criminalizing False Speech Made on Social Networking Sites in A Post-Alvarez, Social Media-Obsessed World’, *Harvard Journal of Law and Technology* 31 (2020), 66–109, at 88–89.

<sup>63</sup> *Police Dep’t of Chi. v. Mosley*, 408 U.S. 92, 95 (1972).

<sup>64</sup> *United States v. Eichman*, 110 S. Ct. 2404, 2410 (1990).

<sup>65</sup> Tompros et al., supra note 62, at 90, referring to *Burson v. Freeman*, 504 U.S. 191, 199 (1992).

<sup>66</sup> See e.g. *United States v. Alvarez*, 567 U.S. 709, 717 (2012). A regulation of unprotected speech may still violate the First Amendment with regard to content discrimination if it includes distinctions among subcategories of speech that cannot be justified. See e.g. *R.A.V. v. City of St. Paul*, 505 U.S. 377 (1992).

<sup>67</sup> *United States v. Williams*, 553 U.S. 285, 322 (dissenting opinion Souter).

<sup>68</sup> *New York Times Co. v. Sullivan*, 376 U.S. 254 (1964), at 279–280. The decision in *Gertz* extended the *NY Times* standard of ‘reckless disregard’ from public officials to public figures and defined these as the persons who, due to their notoriety, achievements, or the rigour of their success, seek the attention of the public. See *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 350 (1974).

<sup>69</sup> 567 U.S. 709 (2012). For a fully analysis of the case, see Tompros et al., supra note 62, at 93–97.

<sup>70</sup> *Ibid.*, at 718.

<sup>71</sup> *Ibid.*, at 710, 726.

<sup>72</sup> N.Y. Penal Law §240.50 reads: ‘A person is guilty of falsely reporting an incident in the third degree when, knowing the information reported, conveyed or circulated to be false or baseless, he or she [...] [i]nitiates or circulates a false report or warning of an alleged occurrence or impending occurrence of a crime, catastrophe or emergency under circumstances in which it is not unlikely that public alarm or inconvenience will result’. Falsely reporting an incident in the third degree is a class A misdemeanor and punishable by up to one year’s imprisonment and a fine of USD 1’000. The statute permits in addition entities providing emergency services to seek restitution for ‘the amount of funds reasonably expended for the purpose of responding’ to false reports.

<sup>73</sup> For a fully-fledged analysis of the law, as well as its possible unconstitutionality post-*Alvarez*, see Tompros et al., supra note 62.

<sup>74</sup> See e.g. Tourkochoriti, supra note 58; Mira Burri, ‘Interfacing Privacy and Trade’, *Case Western Journal of*

speech safeguards under the First Amendment and other domestic statutes,<sup>75</sup> and so insulates American companies from liability.

In the context of online platforms, US law grants an almost perfect safe harbour from liability through Section 230(c)(1) of the Communications Decency Act (CDA),<sup>76</sup> which states that: ‘No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider’. Section 230 grants important substantive and procedural advantages that only enhance the safeguards of the First Amendment.<sup>77</sup> These famous ‘twenty-six words that created the Internet’<sup>78</sup> have been critical in the evolution of online platforms, as well as in their positioning as ‘new governors’ of the online media space.<sup>79</sup> Especially important in the latter context is the possibility that Section 230 enables the intermediary to make a decision in good faith to block or remove content that the intermediary considers ‘to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, *whether or not such material is constitutionally protected*’.<sup>80</sup> Platforms thus have a critical toolkit for content moderation and have used this at times aggressively – for instance, by permanently suspending the Twitter account of President Trump (@realDonaldTrump) on 8 January 2021. Section 230 is, in this sense, a unique type of platforms’ immunity<sup>81</sup> that, next to the strong protection granted under the US Constitution’s First Amendment, shields platforms. In recent years through the channel of free trade agreements, the US has tried to diffuse this unique arrangement, for instance by inserting a similar rule on ‘interactive computer services’ in the updated North American Free Trade Agreement (NAFTA, now called United States Mexico Canada Agreement, USMCA) with Canada and Mexico.<sup>82</sup> Section 230 has not been left without criticism, though, and there have been a number of attempts to constrain its broadly defined immunity, especially in consideration of the changed media environment and the critical role of platforms in it.<sup>83</sup> One attempt that has been successful was the adoption of the 2017 FOSTA<sup>84</sup> to fight online sex trafficking. But there are also other projects in the pipeline aiming at an amendment of Section 230 and addressing the power of Big Tech, such as the ‘Ending Support for Internet Censorship Act’,<sup>85</sup> the ‘Biased

---

*International Law* 53 (2021), 35–88.

<sup>75</sup> Securing the Protection of Our Enduring and Established Constitutional Heritage Act (SPEECH Act), 124 Stat. 2380 (2010). See Eric Goldman, ‘An Overview of the United States’ Section 230 Internet Immunity’, in Giancarlo Frosio (ed), *The Oxford Handbook of Online Intermediary Liability* (Oxford: Oxford University Press, 2020), 155–171.

<sup>76</sup> Communications Decency Act of 1996, (CDA), Pub. L. No. 104-104 (Tit. V), 110 Stat. 133 (8 February 1996), codified at 47 U.S.C. §§223, 230. For a detailed analysis, see Valerie C. Brannon and Eric N. Holmes, ‘Section 230: An Overview’, *Congressional Research Service Report* R46751, 7 April 2021.

<sup>77</sup> Goldman, *supra* note 75; see also Eric Goldman, ‘Why Section 230 Is Better Than the First Amendment’, *Notre Dame Law Review Reflection* 95 (2019), 33–46; also Tanner Bone, ‘How Content Moderation May Expose Social Media Companies to Greater Defamation Liability’, *Washington University Law Review* 98 (2021), 937–963.

<sup>78</sup> Jeff Kloseff, *The Twenty-Six Words That Created the Internet* (Ithaca, NY: Cornell University Press, 2019).

<sup>79</sup> Klonick (2018), *supra* note 27.

<sup>80</sup> 47 U.S.C. §230(c)(2)(A) (emphasis added).

<sup>81</sup> Goldman, *supra* note 75.

<sup>82</sup> See e.g. Mira Burri, ‘Approaches to Digital Trade and Data Flow Regulation across Jurisdictions’, *Journal of World Investment and Trade* 22 (forthcoming 2021).

<sup>83</sup> See e.g. Goldman (2019), *supra* note 77; on the constitutionality of possible Section 230 amendments, see e.g. Brannon and Holmes, *supra* note 77; see also Danielle Keats Citron and Benjamin Wittes, ‘The Internet Will Not Break: Denying Bad Samaritans § 230 Immunity’, *Fordham Law Review* 86 (2017), 401–423 (arguing that platforms should enjoy immunity from liability if they could show that their response to unlawful uses of their services was reasonable).

<sup>84</sup> Allow States and Victims to Fight Online Sex Trafficking Act of 2017, H.R.1865 (115th Cong. 2017-18).

<sup>85</sup> Ending Support for Internet Censorship Act, S. 1914, 116th Cong. (2019).

Algorithm Deterrence Act’,<sup>86</sup> and the ‘Algorithmic Accountability Act’.<sup>87</sup> The laws particularly seek to calibrate the content moderation power of platforms,<sup>88</sup> but their fate is still uncertain. In scholarly debates, there have also been initiatives to use common law doctrines for guidance, such as, in particular, the fiduciary duty-based relationships, such as those between doctors and their patients; lawyers and their clients, and to create new types of obligations for platforms as information fiduciaries<sup>89</sup> in this frame.

### 3. *Developments in the European Union*

#### (i) *Remarks on the General Framework for the Protection of Freedom of Expression*

The developments in Europe with regard to the regulation of platforms have occurred in multiple areas of law and in the broader policy context starting around 2015.<sup>90</sup> While free speech is robustly protected in Europe, the protection does differ from that awarded in the United States and the constitutional balance is differently struck. Importantly in this context, although the right to freedom of expression does not have a direct horizontal application,<sup>91</sup> there is a positive duty of the state to protect it and act as the ‘ultimate guarantor’ of media pluralism.<sup>92</sup> The fundamental right of freedom of expression, both its active and passive dimensions, is enshrined in Article 10 of the European Convention of Human Rights (ECHR)<sup>93</sup>, and the ambit of the protection is wide. The right covers both the content and the form of communication, and applies to any means of dissemination or reception of communication.<sup>94</sup> Much in light of the US Supreme Court’s judgements, the European Court of Human Rights (ECtHR) has held that freedom of expression is protected not only for information and ideas that are favourably accepted, inoffensive or indifferent but also for such that ‘offend, shock or disturb the State or any sector of the population’.<sup>95</sup> Yet, in contrast to the US stance and because of their historical experience, Europeans have not shared the concept of the marketplace of ideas, and different tests have been developed to strike a balance in clashes between different rights, as well as to protect key societal values, such as equality, anti-discrimination and democracy, which may

<sup>86</sup> Biased Algorithm Deterrence Act of 2019, H.R. 492, 116th Cong. (2019).

<sup>87</sup> Algorithmic Accountability Act, S. 1108, 116th Cong. (2019).

<sup>88</sup> For details on and analysis of the legislative proposals, see Bone, *supra* note 77.

<sup>89</sup> The IF model maintains that platforms should be required to abide by fiduciary duties of care, loyalty and confidentiality with regard to their end users. For a discussion, see Jack M. Balkin, ‘Information Fiduciaries and the First Amendment’, *UC Davies Law Review* 49 (2016), 1183–1234; Jack M. Balkin, ‘Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation’, *UC Davies Law Review* 51 (2018), 1149–1210; Jack M. Balkin and Jonathan Zittrain, ‘A Grand Bargain to Make Tech Companies Trustworthy’, *The Atlantic*, 3 October 2016; Lina M. Khan and David E. Pozen, ‘A Skeptical View of Information Fiduciaries’ *Harvard Law Review* 133 (2019), 497–541; Richard S. Whitt, ‘Old School Goes Online: Exploring Fiduciary Obligations of Loyalty and Care in the Digital Platforms Era’, *Santa Clara High Technology Law Review* 36 (2019), 75–131; Claudia E. Haupt, ‘Platforms As Trustees: Information Fiduciaries And The Value Of Analogy’, *Harvard Law Review Forum* 134 (2020), 34–41.

<sup>90</sup> See e.g. Giovanni De Gregorio, ‘Democratising Online Content Moderation: A Constitutional Framework’, *Computer Law and Security Review* 36 (2020), <https://doi.org/10.1016/j.clsr.2019.105374>; Hannah Bloch-Wehba, ‘Global Platform Governance: Private Power in the Shadow of the State’, *SMU Law Review* 73 (2019), 27–80.

<sup>91</sup> *Dink v Turkey* [2010] ECtHR 2668/07, 6102/08, 30079/08, 7072/09 and 7124/09.

<sup>92</sup> *Informationsverein Lentia and Others v Austria* [1993] ECtHR 13914/88; 15041/89; 15717/89; 15779/89; 17207/90.

<sup>93</sup> European Convention on Human Rights, 4 November 1950, 213 U.N.T.S. 221; see also Article 11 (Freedom of Expression and Information) of the Charter of Fundamental Rights of the European Union (CFREU), OJ C [2012] 326/393, 26 October 2012.

<sup>94</sup> See e.g. *Autronic AG v Switzerland* [1990] ECtHR 12726/87; *Schweizerische Radio- und Fernsehgesellschaft SRG v Switzerland* [2012] ECtHR 34124/06.

<sup>95</sup> *Handyside v. United Kingdom* [1976] ECtHR 5493/72, at para. 49.

ultimately lead to constraining freedom of expression.<sup>96</sup> For instance, the ECtHR has permitted in this context that harmful effects on social peace and political stability do justify an interference with freedom of expression.<sup>97</sup> Amongst others, denial of the Holocaust<sup>98</sup> or incitement to hatred or racial discrimination<sup>99</sup> have been held as not protected under Article 10 ECHR. In recent years and because of the critical importance of personal data in the contemporary data-driven society,<sup>100</sup> privacy protection has been given in many cases primacy over the right to freedom of expression. This has been clearly illustrated by the stance of the European Court of Justice (CJEU) in the *Google Spain* case,<sup>101</sup> which famously coined the ‘right to be forgotten’.<sup>102</sup> The Court held that an individual has the right to object to a search engine’s linking to personal information and that the evaluation of such an objection calls for a balancing of rights and interests, in the context of which account must be taken of the significance of the data subject’s rights arising from Articles 7 and 8 of the Charter of Fundamental Rights of the European Union.<sup>103</sup> Effectively, the right to be forgotten, which is now also enshrined under the General Data Protection Regulation (GDPR),<sup>104</sup> trumps, under certain circumstances,<sup>105</sup> the economic freedom of search engines, their freedom of expression, and constrains the passive dimension of the freedom of expression, as it makes information unavailable for the public. Importantly for this chapter’s discussion, the truthfulness of a statement plays an important role when balancing conflicting interests, such as, for instance, hate speech or privacy protection.<sup>106</sup> The Court has also stressed that, while Article 10 affords journalists wide protection, it is not unlimited but ‘subject to the proviso that they are acting in good faith and on an accurate factual basis and provide “reliable and precise” information in

<sup>96</sup> For details, see Oster, *supra* note 49, at Chapter 3; Pollicino, *supra* note 33.

<sup>97</sup> See e.g. *Perinçek v. Switzerland* [2015] ECtHR 27510/08.

<sup>98</sup> See e.g. *Lehideux and Isorni v. France* [1998] ECtHR 55/1997/839/1045; *Garaudy v. France* [2003] ECtHR 65831/01; *Witzsch v. Germany* [2005] ECtHR 7485/03.

<sup>99</sup> See e.g. *Pavel Ivanov v. Russia* [2004] ECtHR 35222/04; *Aksu v. Turkey* [2012] ECtHR 4149/04 and 41029/04.

<sup>100</sup> See e.g. Viktor Mayer-Schönberger and Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work, and Think* (New York: Eamon Dolan/Houghton Mifflin Harcourt, 2013); Bart van der Sloot, Dennis Broeders, and Erik Schrijvers (eds), *Exploring the Boundaries of Big Data* (Amsterdam: Amsterdam University Press, 2016); Burri (2021), *supra* note 74.

<sup>101</sup> Case C-131/12, *Google Spain SL and Google Inc. v. Agencia Española de Protección de Datos (AEPD) and Mario Costeja González*, Judgment of the Court (Grand Chamber) of 13 May 2014, ECR [2014] 317 [hereinafter *Google Spain*].

<sup>102</sup> See e.g. Simon Wechsler, ‘The Right to Remember: The European Convention on Human Rights and the Right to Be Forgotten’, *Columbia Journal of Law and Social Problems* 49 (2015), 135–165; David Hoffman, Paula Bruening and Sophia Carter, ‘The Right to Obscurity: How We Can Implement the Google Spain Decision’, *North Carolina Journal of Law and Technology* 17, (2016), 437–481.

<sup>103</sup> *Google Spain*, at para. 74, referring to Joined Cases C-468/10 and C-469/10, *Asociación Nacional de Establecimientos Financieros de Crédito (ASNEF) and Federación de Comercio Electrónico y Marketing Directo (FECEMD) v. Administración del Estado*, Judgment of 24 November 2011, ECR I-12181, at paras. 38 and 40.

<sup>104</sup> Directive 2016/680 of the European Parliament and of the Council of April 27, 2016, on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection, or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA, OJ L [2016] 119/89. [hereinafter *GDPR*].

<sup>105</sup> *Google Spain*, at para. 88. There is a qualification in para. 99: ‘As the data subject may, in the light of his fundamental rights under Articles 7 and 8 of the Charter, request that the information in question no longer be made available to the general public on account of its inclusion in such a list of results, those rights override, as a rule, not only the economic interest of the operator of the search engine but also the interest of the general public in having access to that information upon a search relating to the data subject’s name. However, that would not be the case if it appeared, for particular reasons, such as the role played by the data subject in public life, that the interference with his fundamental rights is justified by the preponderant interest of the general public in having, on account of its inclusion in the list of results, access to the information in question’.

<sup>106</sup> Oster, *supra* note 49, at 18 and Chapter 3.

accordance with the ethics of journalism',<sup>107</sup> which is to be judged depending on the situation and the technical means used.<sup>108</sup>

The Court has repeatedly emphasized the importance of the Internet for the exercise of freedom of expression, since it provides 'essential tools for participation in activities and discussions concerning political issues and issues of general interest'.<sup>109</sup> Although the Court has not yet had the opportunity to squarely address the duties and responsibilities of social media platforms under Article 10 ECHR,<sup>110</sup> it has tackled cases with regard to automated content moderation.<sup>111</sup> The Court found, in particular, in *Delfi* that a news organization can be held responsible for users' comments if it fails to identify and remove infringing ones, despite the filtering system that the organization had in place.<sup>112</sup> The Court has, however, somewhat moved away from this strict liability for unfiltered comments in *MTE v. Hungary*, where it found that 'this amounts to requiring excessive and impracticable forethought capable of undermining freedom of the right to impart information on the Internet'.<sup>113</sup>

In consideration of the global nature of the Internet, the extraterritorial effect of European precedents appears critical. In the landmark follow-up case to *Google Spain, Google v. CNIL*,<sup>114</sup> which discussed whether the right to be forgotten should have a global reach, the CJEU was cautious and stated that the effect should be limited to the European Union, while using geo-blocking technology in making sure that the right to de-referencing is properly safeguarded.<sup>115</sup> The protection of freedom of expression did play a key role in the Court's decision, as well as the consideration of the different free speech standards around the world,<sup>116</sup> and that Google, if made to comply with a global obligation to de-reference, would effectively be faced with an infringement of the First Amendment in the US. Indeed, as noted earlier, both the First Amendment and Section 230 immunize search engines for their indexing decisions,<sup>117</sup> and operators cannot be legally compelled to implement a right to be forgotten in the United States.<sup>118</sup> Yet, in a later case concerning defamation of an Austrian politician on Facebook, the

<sup>107</sup> *Bladet Tromsø and Stensaas v Norway* [1999] ECtHR 21980/93, at para. 65.

<sup>108</sup> *Stoll v Switzerland* [2007] ECtHR 69698/01.

<sup>109</sup> *Ahmet Yildirim v Turkey* [2012] ECtHR 3111/10, at para. 54.

<sup>110</sup> See Natali Helberger et al., 'A Freedom of Expression Perspective on AI in the Media – with a Special Focus on Editorial Decision Making on Social Media Platforms and in the News Media', *European Journal of Law and Technology* 11 (2020).

<sup>111</sup> See *ibid.*

<sup>112</sup> *Delfi v Estonia* [2015] ECtHR 64569/09.

<sup>113</sup> *MTE v Hungary* [2016] ECtHR 22947/13, at para. 82.

<sup>114</sup> C-507/17, *Google v. Commission nationale de l'informatique et des libertés (CNIL)*, Judgment of 24 September 2019, ECLI:EU:C:2019:772. For a great summary of the case and references to the primary sources, see Columbia Global Freedom of Expression, <https://globalfreedomofexpression.columbia.edu/cases/google-llc-v-national-commission-on-informatics-and-liberty-cnll/>

<sup>115</sup> *Ibid.*, at para. 74: '... the operator is not required to carry out that de-referencing on all versions of its search engine, but on the versions of that search engine corresponding to all the Member States, using, where necessary, measures which, while meeting the legal requirements, effectively prevent or, at the very least, seriously discourage an internet user conducting a search from one of the Member States on the basis of a data subject's name from gaining access via the list of results displayed following that search, to the links which are the subject of that request'.

<sup>116</sup> See *ibid.*, at para. 27; see also Opinion of Advocate General Szpunar delivered on 10 January 2019, ECLI:EU:C:2019:15.

<sup>117</sup> Goldman, *supra* note 75, referring to *Search King, Inc. v. Google Technology, Inc.*, 2003 WL 21464568 (W.D. Okla. 2003); *Langdon v. Google, Inc.*, 474 F. Supp. 2d 622 (D. Del. 2007); *Zhang v. Baidu.com, Inc.*, 10 F. Supp. 3d 433 (S.D.N.Y. 2014); *Google, Inc. v. Hood*, 96 F. Supp. 3d 584 (S.D. Miss. 2015); *e-ventures Worldwide v. Google, Inc.*, 2017 WL 2210029 (M.D. Fla. 2017); *Martin v. Hearst Corporation*, 777 F.3d 546 (2d Cir. 2015).

<sup>118</sup> Goldman, *supra* note 75, referring to *Maughan v. Google Technology, Inc.*, 143 Cal. App. 4th 1242 (Cal. App. Ct. 2006); *Murawski v. Pataki*, 514 F. Supp. 2d 577 (S.D.N.Y. 2007); *Shah v. MyLife.Com, Inc.*, 2012 WL 4863696

CJEU found differently.<sup>119</sup> First and importantly, the Court held that national Courts could order platforms to take down both a specific piece of content and identical content, as well as *equivalent* content (that is, content which conveys the same message, albeit slightly differently worded), as a type of a continued obligation for the platforms. The Court's argument for this extension of the duty was that otherwise, it would be too easy to circumvent an order. However, the Court was careful to delimit the obligation on platforms, by stating that the equivalent information had to be identified in the order, such that the hosting provider did not have to carry out an independent assessment of what constitutes equivalent content. The Court found in addition that such an order was not excessive, since hosting platforms have automated search tools and technologies, and there is no general monitoring obligation involved, which will be contrary to Article 15 of the 2000 E-Commerce Directive<sup>120</sup> that regulates providers' liability for third party content.<sup>121</sup> Second, and critically for the extraterritorial impact of the judgment, the Court found that the E-Commerce Directive does not preclude orders from producing worldwide effects, provided that it is consistent with international law, which is for the Member States to assess. In the first follow-up case after *Glawischnig/Facebook*,<sup>122</sup> the Austrian Supreme Court confirmed and clarified some aspects of the CJEU decision and made it clear that when evaluating the removal of 'information with an equivalent meaning', the balance of interests must not depend on the availability of automated search tools but requires that it can be determined at first glance by a layperson. The Austrian Court also confirmed the possibility of worldwide relief against the infringement of personal rights; noted though that there must be a specific plaintiff's request for this.<sup>123</sup>

Even from this relatively short analysis of the protection of free speech in the United States and in the European Union, it becomes readily evident that the constitutional balances are quite differently struck and the possibilities, indeed the duties, that EU states have to regulate platforms in order to protect vital societal interest are multiple and may lead to certain constraints on freedom of expression, both in its active and passive dimensions. The EU and its Member States also have secondary legislation in place that addresses platforms more squarely. The next section offers a glimpse into these regulatory initiatives.

## (ii) Specific Initiatives Regulating Platforms

The EU framework on platform regulation has evolved over the years and includes different pieces of legislation, some of them predating the predicament of online misinformation. The starting point is the above-mentioned and now relatively dated 2000 E-Commerce Directive, which horizontally regulates the liability of Internet service providers for all types of

---

(D. Or. 2012); *Merritt v. Lexis Nexis*, 2012 WL 6725882 (E.D. Mich. 2012); *Nieman v. Versuslaw, Inc.*, 2012 WL 3201931 (C.D. Ill. 2012); *Getachew v. Google, Inc.*, 491 Fed. Appx. 923 (10th Cir. 2012); *Mmubango v. Google, Inc.*, 2013 WL 664231 (E.D. Pa. 2013); *O'Kroley v. Fastcase Inc.*, 831 F.3d 352 (6th Cir. 2016); *Fakhrian v. Google Inc.*, 2016 WL 1650705 (Cal. App. Ct. 2016); *Despot v. Baltimore Life Insurance Co.*, 2016 WL 4148085 (W.D. Pa. 2016); *Manchanda v. Google, Inc.*, 2016 WL 6806250 (S.D.N.Y. 2016).

<sup>119</sup> Case C-18/18 *Eva Glawischnig-Piesczek v Facebook Ireland Limited*, Judgment of 3 October 2019, ECLI:EU:C:2019:821. For a great summary of the case and references to the primary sources, see Columbia Global Freedom of Expression, <https://globalfreedomofexpression.columbia.edu/cases/glawischnig-piesczek-v-facebook-ireland-limited/>; for a critique of the case, see Daphne Keller, 'Dolphins in the Net: Internet Content Filters and the Advocate General's *Glawischnig-Piesczek v. Facebook Ireland* Opinion', Stanford Center for Internet and Society, 4 September 2019.

<sup>120</sup> 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 (Directive on electronic commerce), OJ L [2000] 178/1.

<sup>121</sup> See in particular Article 14 E-Commerce Directive.

<sup>122</sup> Austrian Supreme Court, *ORF/Facebook*, Judgment 4Ob36/20b of 30 March 2020.

<sup>123</sup> There was not such a request involved in the case at issue.

infringement (excluding tax and data protection, audiovisual services and gambling). The E-Commerce Directive specifies, in particular, the ‘notice and takedown’ regime depending on the different types of interaction between the platform and the content,<sup>124</sup> while prohibiting general monitoring obligations as a guarantee for fundamental rights protection.<sup>125</sup> The E-Commerce Directive also includes specific norms on co- and self-regulatory measures<sup>126</sup> and a duty to cooperate with competent authorities.<sup>127</sup> In 2018 the European Commission provided more specific guidance with the Recommendation to fight illegal content online,<sup>128</sup> which set out the general principles for all types of illegal content online and recommended stricter moderation for terrorist content. In particular, the Recommendation addressed the ‘notice and takedown’ procedures and specified that they must (i) be effective, precise and adequately substantiated; (ii) respect the rights of content providers with possibilities of counter-notices and out-of-court dispute settlement and (iii) be transparent.<sup>129</sup> With regard to the proactive measures that operators should adopt, the Recommendation encouraged appropriate, proportionate and specific measures, which may use automated means but with safeguards in place, in particular human oversight and verification.<sup>130</sup> The Recommendation also encouraged closer cooperation with national judicial and administrative authorities as well as with trusted flaggers with the necessary expertise and determined on a clear and objective basis; it fostered cooperation among hosting providers, in particular smaller ones that may have less capacity to tackle illegal content.<sup>131</sup>

In addition to the intermediaries’ liability regulation, the EU has had since 1989 an instrument to regulate media services,<sup>132</sup> whose scope of application has been extended to cover online content services as well over the years and in light of technological developments. The most recent revision of the Audiovisual Media Service Directive (AVMSD)<sup>133</sup> of 2018 also applies to the so-called video-sharing platforms (VSPs), which in essence address user-generated

<sup>124</sup> Articles 12–14 E-Commerce Directive.

<sup>125</sup> Article 15 E-Commerce Directive. In *Scarlet v SABAM* (Case C-70/10, *Scarlet Extended SA v Société belge des auteurs, compositeurs et éditeurs SCRL (SABAM)*, ECLI:EU:C:2011:771), the Belgian collecting society SABAM applied for a permanent order requiring a network access provider to monitor and block peer-to-peer transmission of music files from SABAM’s catalogue. The CJEU decided that a broad order of the type requested would go both against the prohibition of general monitoring obligations of the E-Commerce Directive and the fundamental rights of Internet users to the protection of their personal data and freedom of expression guaranteed under the EU Charter of Fundamental Rights. See also Case C-360/10, *Belgische Vereniging van Auteurs, Componisten en Uitgevers CVBA (SABAM) v Netlog NV*, ECLI:EU:C:2012:85. Specific monitoring obligations have been however found not in violation of Article 15 E-Commerce Directive (see Alexandre De Stree, Aleksandra Kuczerawy and Michèle Ledger, ‘Online Platforms and Services’, in Laurent Garzaniti et al. (eds), *Electronic Communications, Audiovisual Services and the Internet* (London: Sweet and Maxwell, 2021), 125–157).

<sup>126</sup> Article 16 E-Commerce Directive

<sup>127</sup> Article 15(2) E-Commerce Directive.

<sup>128</sup> European Commission, Recommendation 2018/334 of 1 March 2018 on measures to effectively tackle illegal content online, OJ L [2018] 63/50. See also European Commission, Tackling Illegal Content Online. Towards an Enhanced Responsibility for Online Platforms, COM (2017) 555 final, 28 September 2017.

<sup>129</sup> Recommendation 2018/334, *ibid.*, at points 5–17.

<sup>130</sup> Recommendation 2018/334, *ibid.*, at points 16–21.

<sup>131</sup> Recommendation 2018/334, *ibid.*, at points 22–28. For an evaluation of the rules, see Alexandre de Stree et al., *Online Platforms’ Moderation of Illegal Content Online: Law, Practices and Options for Reform*, Study for the European Parliament (Brussels: European Parliament, 2020), at 22–23.

<sup>132</sup> Council Directive 89/552/EEC of 3 October 1989 on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the pursuit of television broadcasting activities, OJ L [1989] 298/23.

<sup>133</sup> Directive 2010/13 of the European Parliament and of the Council of 10 March 2010 on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audio-visual media services (Audio-Visual Media Services Directive), OJ L [2010] 95/1, as amended by Directive 2018/1808, OJ L [2018] 303/69 [hereinafter AVMSD].

content platforms.<sup>134</sup> VSPs face lighter duties and responsibilities than broadcasting channels and platforms with editorial responsibility (such as Netflix, for instance) but must take appropriate measures to protect the general public from certain types of online content, in particular racism and xenophobia, as well as hate speech based upon the illegal grounds as laid down in the EU Charter of Fundamental Rights (sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation).<sup>135</sup> VSPs must also protect minors from content, which may impair their physical, mental or moral development.<sup>136</sup> The AVMSD notes that the measures must be ‘appropriate’ in the light of the nature of the content, the potential harm, the characteristics of the category of persons to be protected, the rights and legitimate interests at stake, as well as ‘practicable and proportionate’, taking into account the size of the VSP and the nature of the provided service.<sup>137</sup> In addition to the AVMSD, platforms bear certain distinct duties with regard to fighting terrorist content online under the Counter-Terrorism Directive<sup>138</sup> and the Directive against Child Sexual Abuse.<sup>139</sup>

With regard to platforms, the EU has also adopted a number of more recent measures of self- and co-regulatory nature. Noteworthy in the context of this chapter’s discussion is first, the 2016 Code of Conduct on Countering Illegal Hate Speech Online<sup>140</sup> and second, the 2018 Code of Practice on Disinformation.<sup>141</sup> Under the Illegal Hate Speech Code, platforms have made the following, amongst other, commitments: (i) to draw users’ attention to the types of content not allowed by their Community Standards/Guidelines and specify that they prohibit the promotion of incitement to violence and hateful behaviour; (ii) to put in place clear, effective and speedy processes to review notifications of illegal hate speech and remove illegal content; (iii) to encourage the reporting of illegal hate speech by experts, including through partnerships with civil society organizations; (iv) to strengthen communication and cooperation between the online platforms and the national authorities, in particular with regard to procedures for

---

<sup>134</sup> Article 1(1aa) AVMSD defines the ‘video-sharing platform service’ as ‘a service as defined by Articles 56 and 57 TFEU, where the principal purpose of the service or of a dissociable section thereof or an essential functionality of the service is devoted to providing programmes, user-generated videos, or both, to the general public, for which the video-sharing platform provider does not have editorial responsibility, in order to inform, entertain or educate, by means of electronic communications networks [...] and the organisation of which is determined by the video-sharing platform provider, including by automatic means or algorithms in particular by displaying, tagging and sequencing’. See also European Commission, Guidelines on the practical application of the essential functionality criterion of the definition of a ‘video-sharing platform service’ under the Audiovisual Media Services Directive, OJ C [2020] 223/3, 7 July 2020.

<sup>135</sup> Article 28b (1b) and (1c) AVMSD.

<sup>136</sup> Article 28b (1a) AVMSD.

<sup>137</sup> Article 28b (3) AVMSD. The AVMSD lists certain appropriate measures, such as transparent and user-friendly mechanisms to report and flag the content and easy- to-use and effective procedures for the handling and resolution of users’ complaints.

<sup>138</sup> Directive 2017/541 of the European Parliament and of the Council of 15 March 2017 on combating terrorism and replacing Council Framework Decision 2002/475/JHA and amending Council Decision 2005/671/JHA, OJ L [2017] 88/6.

<sup>139</sup> Directive 2011/93/EU of the European Parliament and of the Council of 13 December 2011 on combating the sexual abuse and sexual exploitation of children and child pornography, and replacing Council Framework Decision 2004/68/JHA, OJ L [2011] 335/1. For an analysis of both documents, see e.g. de Streele et al., *supra* note 131, at 25–29.

<sup>140</sup> The Code was signed in 2016 by Facebook, Microsoft, Twitter and YouTube. Google+, Instagram, Dailymotion and Snapchat and Jeuxvideo.com joined subsequently. The Code’s text is available at: [https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online\\_en](https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en)

<sup>141</sup> The Code was signed by Facebook, Google and Twitter, Mozilla, as well as by advertisers and parts of the advertising industry in October 2018; Microsoft joined in May 2019, while TikTok became a signatory in June 2020. Code’s text is available at: <https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation>.

submitting notifications. The Code of Practice on Disinformation involves similar obligations with regard to transparency, due processes and cooperation. It seeks, in particular, to ensure the credibility of information and improve content moderation practices, for instance, by closing false accounts; removing bots; investing in technologies that help users make informed decisions when receiving false information (e.g. through reliability indicators/trust markers); prioritizing relevant and authentic information; facilitating the finding of alternative content on issues of general interest; improving transparency of political and issue-based advertising; empowering the research community, fact-checkers and other relevant stakeholders (e.g. with better access to data).<sup>142</sup> In May 2021, very much in light of the Covid-19 infodemic, the European Commission presented a new Guidance to strengthen the Code of Practice on Disinformation,<sup>143</sup> which aims to address gaps and shortcomings<sup>144</sup> and create a more transparent, safe and trustworthy online environment. The Guidance also aims at evolving the existing Code of Practice towards a co-regulatory instrument foreseen under the Digital Services Act (DSA),<sup>145</sup> which would increase the bindingness of the rules and the level of scrutiny by the European Commission. Building upon the newly announced European Democracy Action Plan (EDAP),<sup>146</sup> the DSA, which would replace the 2000 E-Commerce Directive, introduces wide-ranging transparency measures around content moderation and advertising, and proposes binding and enforceable legal obligations, in particular for very large online platforms,<sup>147</sup> to assess and address systemic risks for fundamental rights or presented by the intentional manipulation of their services. When adopted, the DSA would create a new basis for the regulation of platforms and substantially increase their duties and responsibilities, also with regard to fighting misinformation online.<sup>148</sup> The DSA would also harmonize the rules across different EU Member States' jurisdictions in this context, which have so far developed independently.<sup>149</sup>

One EU Member State's law that has been particularly controversial in policy and academic circles on both sides of the Atlantic<sup>150</sup> has been Germany's Network Enforcement Act

---

<sup>142</sup> The Code includes also an annex identifying best practices that signatories will apply to implement the Code's commitments. For all documents, see <https://ec.europa.eu/digital-single-market/en/news/code-practice-disinformation>

<sup>143</sup> European Commission, Guidance on Strengthening the Code of Practice on Disinformation Brussels, COM(2021) 262 final, 26 May 2021.

<sup>144</sup> The Commission's Assessment of the Code of Practice in 2020 revealed in particular include inconsistent and incomplete application of the Code across platforms and Member States, limitations intrinsic to the self-regulatory nature of the Code, as well as gaps in the coverage of the Code's commitments. The assessment also highlighted the lack of an appropriate monitoring mechanism, including key performance indicators, lack of commitments on access to platforms' data for research on disinformation and limited participation from stakeholders, in particular from the advertising sector. See European Commission (2021), *ibid.* and European Commission, Assessment of the Code of Practice on Disinformation: Achievements and Areas for Further Improvement, SWD(2020)180, 10 September 2020.

<sup>145</sup> Proposal for a Regulation of the European Parliament and the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC, COM(2020) 825 final, 15 December 2020.

<sup>146</sup> European Commission Communication on the European Democracy Action Plan, COM(2020) 790 final, 3 December 2020.

<sup>147</sup> The DSA defines very large platforms in Article 25 as online platforms which provide their services to a number of average monthly active recipients of the service in the EU corresponding to 10% of the EU's population.

<sup>148</sup> For an overview of the new obligations depending on the type of platform, see [https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment\\_en](https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment_en)

<sup>149</sup> For an overview of the different initiatives, see e.g. de Stree et al., *supra* note 131; Roudik et al., *supra* note 56.

<sup>150</sup> See e.g. David Kaye, 'How Europe's New Internet Laws Threaten Freedom of Expression Recent Regulations Risk Censoring Legitimate Content', *Foreign Affairs*, 18 December 2017; Claudia E. Haupt, 'Regulating Speech Online: Free Speech Values in Constitutional Frames', *Washington University Law Review* 99 (forthcoming 2021); see also Patrick Zurth, 'The German NetzDG as Role Model or Cautionary Tale? Implications for the Debate on

(Netzwerkdurchsetzungsgesetz, NetzDG) enacted in 2017 and also known as the ‘Facebook Act’.<sup>151</sup> While the NetzDG does not criminalize new activities, it seeks to enable better enforcement of the German Criminal Code online. For this purpose, the NetzDG enumerates discrete criminal code provisions whose violation it sanctions in the digital space. These include, amongst others, dissemination of propaganda material or use of symbols of unconstitutional organizations; encouragement of the commission of a serious violent offence endangering the state; commission of treasonous forgery; public incitement to crime; incitement to hatred, and defamation.<sup>152</sup> The NetzDG is limited in its scope and only applicable to social media networks that have two million or more registered users in Germany. Social media networks are defined as ‘telemedia service providers that operate online platforms with the intent to make a profit and on which users can share content with other users or make that content publicly available’.<sup>153</sup> The NetzDG requires social media networks to remove ‘manifestly illegal’ content within 24 hours after the content is flagged; other illegal content must be removed within 7 days after receiving a notification.<sup>154</sup> The social media platforms are also obliged to offer their users an easy and transparent complaint mechanism that is constantly available; the decisions taken with regard to the complaint and the reasoning behind accepting or rejecting it must, without delay, be communicated to both the complainant and the affected user.<sup>155</sup> In addition, the NetzDG includes reporting requirements for platforms that receive more than 100 complaints of unlawful postings per calendar year.<sup>156</sup> A social media network that intentionally or negligently violates certain of its duties under the NetzDG may be fined up to EUR 50 million.<sup>157</sup>

Despite being highly controversial, in particular as to its constitutionality,<sup>158</sup> and as to creating real concerns for freedom of speech because of outsourcing of legal enforcement to private entities and over-blocking,<sup>159</sup> there have been efforts to update the NetzDG. The revision

---

Social Media Liability’, *Fordham Intellectual Property, Media and Entertainment Law Journal* 31 (2021), 1084–1153; Heidi Tworek and Paddy Leerssen, ‘An Analysis of Germany’s NetzDG Law’, *Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression* (2019); Danielle Keats Citron, ‘Extremist Speech, Compelled Conformity, and Censorship Creep’, *Notre Dame Law Review* 93 (2018), 1035–1072.

<sup>151</sup> Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken [NetzDG] [Network Enforcement Act], 1 September 2017 [BGBl I] at 3352. The law entered into force on 1 January 2018.

<sup>152</sup> § 1(3) NetzDG, referring to §§ 86, 86a, 89a, 91, 100a, 111, 126, 129 bis 129b, 130, 131, 140, 166, 184b, 185 bis 187, 201a, 241 and 269 of the German Criminal Code.

<sup>153</sup> § 1(1) NetzDG. Platforms that post original journalistic content, email or messaging services are not covered.

<sup>154</sup> The deadline may be extended if additional facts are necessary to determine the truthfulness of the information or if the social network hires an outside agency to perform the vetting process.

<sup>155</sup> § 3 paras. 1 and 2 NetzDG.

<sup>156</sup> § 2 paras. 1 and 2 NetzDG. The report has to be published in German in the Federal Gazette and on the website of the social media network one month after the end of each half-year period. The report must be easily identifiable, immediately accessible, and permanently available. It must include information on the general efforts to prevent illegal actions on the platform, a description of the complaint procedure, the number of complaints received, the number and qualifications of employees who are handling the complaints, the network’s association memberships, the number of times an external party has been used to decide the illegality of the content, the number of complaints that led to the content being deleted, the time it took to delete the content, and measures that were taken to inform the complainant and the member who posted the deleted content.

<sup>157</sup> § 4 NetzDG, in conjunction with Gesetz über Ordnungswidrigkeiten [OWiG] [Act on Regulatory Offenses], 19 February 1987 [BGBl. I] at 602, as amended, § 30(2). The fine is rendered by the Department of Justice upon a Court decision. The decision of the Court is final and binding on the Department of Justice.

<sup>158</sup> See e.g. Haupt, supra note 150; also Nikolaus Guggenberger, ‘Das Netzwerkdurchsetzungsgesetz – schön gedacht, schlecht gemacht’, *Zeitschrift für Rechtspolitik* (2017), 98–101; Georg Nolte, ‘Hate-Speech, Fake-News, das Netzwerkdurchsetzungsgesetz und Vielfaltsicherung durch Suchmaschinen’, *Zeitschrift für Urheber- und Medienrecht* 61 (2017), 552–565.

<sup>159</sup> See e.g. Zurth, supra note 150; see also the references listed in note 150 above.

projects have been in line with other legislative initiatives in Germany introduced by the national implementation of the AVMSD and the new law against right-wing extremism and hate crimes.<sup>160</sup> The updated NetzDG<sup>161</sup> has been adopted in June 2021. It now also covers VSPs,<sup>162</sup> as well as includes a number of new rules with regard to the simplification of the reporting channels for the complaint procedure and additional information obligations for the half-yearly transparency reports of the platform operators (for instance, with regard to the so-called ‘put backs’). In light of the over-blocking concerns, the amendment of the NetzDG provides that users may request a review of the platform provider’s decision to remove or retain reported content and have a right to have the content restored.<sup>163</sup>

### C. Appraisal of the Emergent Regulatory Framework for Platforms

The Covid-19 crisis has recently and painfully revealed the perils of online disinformation, and spurred a flurry of initiatives around the world to curb this negative phenomenon and regulate information platforms to this effect. As Goldsmith and Keane Woods have pointed out, hopefully sometime in the near future, this grave period for global societies will be over, but ‘when the crisis is gone, there is no unregulated “normal” to return to’,<sup>164</sup> as the general trend towards the regulation of digital speech will not abate.<sup>165</sup> This chapter has offered some insights into the emergent regulatory framework for information platforms and revealed the different approaches across jurisdictions with a distinct focus on the United States and the European Union. It became evident that the different constitutional traditions and understandings of the role of the State in the protection of fundamental rights have led to the emergence of very different regulatory environments. Both come with certain pros and cons and raise questions as to the proper balance between freedom of speech and the protection of other vital public interests and individual rights. One should acknowledge in this context in particular that social media platforms, despite their essential role in contemporary communicative processes, are still relatively young, and the risk of over-regulation and unintended consequences is real.<sup>166</sup> In the EU context, while the regulatory efforts are properly based on the international and constitutional frameworks, it remains critical that there are sufficient substantive and procedural safeguards in place that take the different interests involved into account and recognize that with digital speech, we do face a ‘problem of many hands’, where there is a corresponding need to conceptualize a framework with participation of and different responsibilities for all stakeholders – platforms, users, civil society and governments.<sup>167</sup> Otherwise, many of the already recognized dangers associated with speech regulation in undemocratic societies may

---

<sup>160</sup> Gesetz zur Bekämpfung des Rechtsextremismus und der Hasskriminalität, 30 March 2021 [BGBl. I], at 441.

<sup>161</sup> Gesetz zur Änderung des Netzwerkdurchsetzungsgesetzes, 3 June 2021 [BGBl. I] at 1436.

<sup>162</sup> § 3d amended NetzDG.

<sup>163</sup> For overview of the changes, see

<https://www.bmju.de/SharedDocs/Gesetzgebungsverfahren/DE/NetzDGAendG.html> (in German).

<sup>164</sup> Jack Goldsmith and Andrew Keane Woods, ‘Internet Speech Will Never Go Back to Normal’, *The Atlantic*, 26 April 2020.

<sup>165</sup> *Ibid.*

<sup>166</sup> See e.g. Ashutosh Bhagwat, ‘The Law of Facebook’, *UC Davies Law Review* 54 (2021), 2353–2403.

<sup>167</sup> Natali Helberger, Jo Pierson and Thomas Poell, ‘Governing Online Platforms: From Contested to Cooperative Responsibility’, *The Information Society* (2017), <https://doi.org/10.1080/01972243.2017.1391913>, at 2; see also Michele Finck, ‘Digital Co-regulation: Designing a Supranational Legal Framework for the Platform Economy’, *European Law Review* 43 (2018), 47–68 (arguing for a co-regulatory model of platform regulation); Florian Saurwein and Charlotte Spencer-Smith, ‘Combating Disinformation on Social Media: Multilevel Governance and Distributed Accountability in Europe’, *Digital Journalism* 8 (2020), 820–841 (mapping the different governance approaches).

occur,<sup>168</sup> and digital innovation may be seriously hindered. As for the United States, it remains uncertain whether and in what form platform liability would be reformed against the backdrop of the high standards of protection of the First Amendment, and to what extent platforms themselves would use their leeway and more aggressively moderate content. At this point in time, we do have ‘geographically-segmented speech’,<sup>169</sup> and regulators would need to make sure that this situation is not exacerbated, either by means of geographic filtering and geo-blocking or by agreeing upon new baselines of transparency and accountability, and enhanced cooperation between governments and tech companies.<sup>170</sup>

## Bibliography

- Darrin Baines and Robert J. R. Elliott, ‘Defining Misinformation, Disinformation and Malinformation: An Urgent Need for Clarity during the COVID-19 Infodemic’, *University of Birmingham Department of Economics Discussion Papers 20-06* (2020)
- Jack M. Balkin, ‘Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation’, *University of California Davis Law Review* 51 (2018), 1149–1210
- Jack M. Balkin, ‘Information Fiduciaries and the First Amendment’, *UC Davies Law Review* 49 (2016), 1183–1234
- Jack M. Balkin, ‘Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation’, *UC Davies Law Review* 51 (2018), 1149–1210
- Jack M. Balkin, ‘Free Speech Is a Triangle’, *Columbia Law Review* 118 (2018), 2011–2055
- Jack M. Balkin and Jonathan Zittrain, ‘A Grand Bargain to Make Tech Companies Trustworthy’, *The Atlantic*, 3 October 2016
- Kenneth A. Bamberger and Orly Lobel, ‘Platform Market Power’, *Berkeley Technology Law Journal* 32 (2018) 32, 1052–1092
- Judith Bayer et al., *Disinformation and Propaganda – Impact on the Functioning of the Rule of Law in the EU and its Member States*, Study for the European Parliament (Brussels: European Parliament, 2019)
- Yochai Benkler, *The Wealth of Networks: How Social Production Transforms Markets and Freedom* (New Haven, CT: Yale University Press, 2006)
- Yochai Benkler et al., *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics* (Oxford: Oxford University Press, 2018)
- Ashutosh Bhagwat, ‘The Law of Facebook’, *UC Davies Law Review* 54 (2021), 2353–2403.
- Hannah Bloch-Wehba, ‘Global Platform Governance: Private Power in the Shadow of the State’, *SMU Law Review* 73 (2019), 27–80
- Tanner Bone, ‘How Content Moderation May Expose Social Media Companies to Greater Defamation Liability’, *Washington University Law Review* 98 (2021), 937–963
- Kalina Bontcheva and Julie Posetti (eds), *Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression*, Broadband Commission research report on ‘Freedom of Expression and Addressing Disinformation on the Internet’ (Geneva/Paris: ITU/UNESCO, 2020)
- Engin Bozdogan, ‘Bias in Algorithmic Filtering and Personalization’, *Ethics and Information Technology* 15 (2013), 209–227
- Valerie C. Brannon and Eric N. Holmes, ‘Section 230: An Overview’, *Congressional Research Service Report R46751*, 7 April 2021

<sup>168</sup> See e.g. with regard to Russia, Evgeni Moyakina and Alexander Tabachnik, ‘Struggling to Strike the Right Balance between Interests at Stake: The “Yarovaya”, “Fake news” and “Disrespect” Laws as Examples of Ill-Conceived Legislation in the Age of Modern Rechnology’, *Computer Law and Security Review* 40 (2021), <https://doi.org/10.1016/j.clsr.2020.105512>; for a discussion of other cases, see Roudik et al., supra note 56.

<sup>169</sup> Jennifer Daskal, ‘Speech Across Borders’, *Virginia Law Review* 105 (2019), 1605–1666, at 1605.

<sup>170</sup> Ibid.

- J. Scott Brennen et al., *Types, Sources, and Claims of COVID-19 Misinformation* (Oxford: Reuters Institute for the Study of Journalism, 2020)
- Mira Burri, *Public Service Broadcasting 3.0: Legal Design for the Digital Present* (London: Routledge, 2015)
- Mira Burri, 'Understanding the Implications of Big Data and Big Data Analytics for Competition Law: An Attempt for a Primer', in Klaus Mathis and Avishalom Tor (eds), *New Developments in Competition Behavioural Law and Economics* (Berlin: Springer, 2019), 241–263
- Mira Burri, 'Discoverability of Local, National and Regional Content Online', *A Thought Leadership Paper written for the Canadian Commission for UNESCO and Canadian Heritage*, 7 February 2019
- Mira Burri, 'Interfacing Privacy and Trade', *Case Western Journal of International Law* 53 (2021), 35–88
- Mira Burri, 'Approaches to Digital Trade and Data Flow Regulation across Jurisdictions', *Journal of World Investment and Trade* 22 (forthcoming 2021)
- Mira Burri and Zaira Zihlmann, 'Intermediaries' Liability in Light of the Recent EU Copyright Reform', *Indian Journal of Intellectual Property Law* 11 (2021)
- Center for Information Technology and Society – UC Santa Barbara, *A Brief History of Fake News* (Santa Barbara, CA: Center for Information Technology and Society, 2019).
- Center for Countering Digital Hate, *Malgorithm: How Instagram's Algorithm Publishes Misinformation and Hate to Millions during a Pandemic* (London and Washington, DC: Center for Countering Digital Hate, 2020)
- Anupam Chander, 'How Law Made Silicon Valley', *Emory Law Journal* 63 (2014), 639–694
- Bobby Chesney and Danielle Citron, 'Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security', *California Law Review* 107 (2019), 1753–1820
- Danielle Keats Citron, 'Extremist Speech, Compelled Conformity, and Censorship Creep', *Notre Dame Law Review* 93 (2018), 1035–1072
- Danielle Keats Citron and Benjamin Wittes, 'The Internet Will Not Break: Denying Bad Samaritans § 230 Immunity', *Fordham Law Review* 86 (2017), 401–423
- Julie E. Cohen, 'Law for the Platform Economy', *University of California Davis Law Review* 51 (2018), 133–204
- Peter Dahlgren, 'The Internet, Public Spheres, and Political Communication', *Political Communication* 22 (2005), 147–162
- Jennifer Daskal, 'The Un-Territoriality of Data', *The Yale Law Journal* 125 (2016), 326–398
- Jennifer Daskal, 'Speech Across Borders', *Virginia Law Review* 105 (2019), 1605–1666
- Giovanni De Gregorio, 'Democratising Online Content Moderation: A Constitutional Framework', *Computer Law and Security Review* 36 (2020), <https://doi.org/10.1016/j.clsr.2019.105374>
- Laura DeNardis, *Protocol Politics: The Globalization of Internet Governance* (Cambridge: MIT Press, 2009)
- Alexandre De Stree, Aleksandra Kuczerawy and Michèle Ledger, 'Online Platforms and Services', in Laurent Garzaniti et al. (eds), *Electronic Communications, Audiovisual Services and the Internet* (London: Sweet and Maxwell, 2021), 125–157
- Kristin E. Eichensehr, 'The Cyber-Law of Nations', *The Georgetown Law Journal* 103 (2015), 317–380
- Kristin E. Eichensehr, 'Data Extraterritoriality', *Texas Law Review* 95 (2017), 145–160
- Adam M. Enders, 'The Different Forms of COVID-19 Misinformation and Their Consequences', *The Harvard Kennedy School (HKS) Misinformation Review* (2020)
- European Data Protection Supervisor (EDPS), *Privacy and Competitiveness in the Age of Big Data*, 2014
- European Commission, Online Platforms and the Digital Single Market, COM(2016) 288 final, 25 May 2016

- European Commission, Assessment of the Code of Practice on Disinformation: Achievements and Areas for Further Improvement, SWD(2020)180, 10 September 2020
- European Commission, Communication on the European Democracy Action Plan, COM(2020) 790 final, 3 December 2020
- European Commission, Proposal for a Regulation of the European Parliament and the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC, COM(2020) 825 final, 15 December 2020
- European Commission, Guidance on Strengthening the Code of Practice on Disinformation Brussels, COM(2021) 262 final, 26 May 2021
- European Commission and the High Representative of the Union for Foreign Affairs and Security Policy, Tackling COVID-19 Disinformation: Getting the Facts Right, JOIN(2020) 8 final, 10 June 2020
- Ariel Ezrachi and Maurice E. Stucke, *Virtual Competition: The Promise and Perils of the Algorithm-Driven Economy* (Cambridge, MA: Harvard University Press, 2016)
- Daniel A. Farber, *The First Amendment: Concepts and Insights* (St. Paul, MN: Foundation Press, 2019)
- Noah R. Feldman and Kathleen M. Sullivan, *First Amendment Law*, 7th edn. (St. Paul, MN: West Academic, 2019)
- Michele Finck, 'Digital Co-regulation: Designing a Supranational Legal Framework for the Platform Economy', *European Law Review* 43 (2018), 47–68
- Brett M. Frishmann, *Infrastructure: The Social Value of Shared Resources* (Oxford: Oxford University Press, 2012)
- Giancarlo Frosio (ed), *Oxford Handbook of Online Intermediary Liability* (Oxford: Oxford University Press, 2020)
- Jack Goldsmith and Tim Wu, *Who Controls the Internet? Illusions of a Borderless World* (Oxford: Oxford University Press, 2001)
- Jack Goldsmith and Andrew Keane Woods, 'Internet Speech Will Never Go Back to Normal', *The Atlantic*, 26 April 2020
- Eric Goldman, 'Why Section 230 Is Better Than the First Amendment', *Notre Dame Law Review Reflection* 95 (2019), 33–46
- Eric Goldman, 'An Overview of the United States' Section 230 Internet Immunity', in Giancarlo Frosio (ed), *The Oxford Handbook of Online Intermediary Liability* (Oxford: Oxford University Press, 2020), 155–171
- Ellen P. Goodman, 'Media Policy Out of the Box: Content Abundance, Attention Scarcity, and the Failures of Digital Markets', *Berkeley Technology Law Journal* 19 (2004), 1389–1472
- Marc H. Greenberg, 'A Return to Lilliput: The LICRA v. Yahoo – Case and the Regulation of Online Content in the World Market', *Berkeley Technology Law Review* 18 (2003), 1191–1258
- Nikolaus Guggenberger, 'Das Netzwerkdurchsetzungsgesetz – schön gedacht, schlecht gemacht', *Zeitschrift für Rechtspolitik* (2017), 98–101
- Cherilyn Ireton and Julie Posetti (eds), *Journalism, 'Fake News' and Disinformation* (Paris: UNESCO, 2018)
- Claudia E. Haupt, 'Regulating Hate Speech: Damned If You Do and Damned If You Don't – Lessons Learned from Comparing the German and U.S. Approaches', *Boston University International Law Journal* 23 (2005), 300–335
- Claudia E. Haupt, 'Platforms as Trustees: Information Fiduciaries And The Value Of Analogy', *Harvard Law Review Forum* 134 (2020), 34–41
- Claudia E. Haupt, 'Regulating Speech Online: Free Speech Values in Constitutional Frames', *Washington University Law Review* 99 (forthcoming 2021)
- Natali Helberger, 'Diversity by Design', *Journal of Information Policy* 1 (2011), 441–469
- Natali Helberger, 'Exposure Diversity as a Policy Goal', *Journal of Media Law* 4 (2012), 65–92
- Natali Helberger, Katharina Kleinen-von Königlöw and Rob van der Noll, 'Regulating the New

- Information Intermediaries as Gatekeepers of Information Diversity', *info* 6 (2015), 50–71
- Natali Helberger, Jo Pierson and Thomas Poell, 'Governing Online Platforms: From Contested to Cooperative Responsibility', *The Information Society* (2017), <https://doi.org/10.1080/01972243.2017.1391913>
- Natali Helberger et al., 'A Freedom of Expression Perspective on AI in the Media – with a Special Focus on Editorial Decision Making on Social Media Platforms and in the News Media', *European Journal of Law and Technology* 11 (2020)
- Christian P. Hoffman et al. 'Diversity by Choice: Applying a Social Cognitive Perspective to the Role of Public Service Media in the Digital Age', *International Journal of Communication* 9 (2015), 1360–1381
- David Hoffman, Paula Bruening and Sophia Carter, 'The Right to Obscurity: How We Can Implement the Google Spain Decision', *North Carolina Journal of Law and Technology* 17, (2016), 437–481
- David Kaye, 'How Europe's New Internet Laws Threaten Freedom of Expression Recent Regulations Risk Censoring Legitimate Content', *Foreign Affairs*, 18 December 2017
- David Kaye, *Speech Policy: The Struggle to Govern the Internet* (New York: Columbia Global Reports, 2019)
- Daphne Keller, 'Dolphins in the Net: Internet Content Filters and the Advocate General's *Glawischnig-Piesczek v. Facebook Ireland* Opinion', Stanford Center for Internet and Society, 4 September 2019
- Paul Keller, *European and International Media Law: Liberal Democracy, Trade, and the New Media* (Oxford: Oxford University Press, 2011)
- Lina M. Khan and David E. Pozen, 'A Skeptical View of Information Fiduciaries', *Harvard Law Review* 133 (2019), 497–541
- Rasmus Kleis Nielsen et al., *Communications in the Coronavirus Crisis: Lessons for the Second Wave* (Oxford: Reuters Institute for the Study of Journalism, 2020)
- Kate Klonick, 'The New Governors: The People, Rules, and Processes Governing Online Speech', *Harvard Law Review* 131 (2018), 1598–1670
- Kate Klonick, 'The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression', *The Yale Law Journal* 129 (2020), 2418–2299
- Jeff Kloseff, *The Twenty-Six Words That Created the Internet* (Ithaca, NY: Cornell University Press, 2019)
- Aleksi Knuutila et al., Covid-Related Misinformation on YouTube: The Spread of Misinformation Videos on Social Media and the Effectiveness of Platform Policies, Oxford Internet Institute, COMPROP Data Memo 2020.6, 21 September 2020
- Daniel Kreiss and Shannon C. McGregor, 'Technology Firms Shape Political Communication: The Work of Microsoft, Facebook, Twitter, and Google with Campaigns During the 2016 US Presidential Cycle', *Political Communication* (2017), 1–23
- Michael Latzer et al. 'The Economics of Algorithmic Selection on the Internet', *Media Change and Innovation Working Paper* (2014)
- Lawrence Lessig, *Remix: Making Art and Commerce Thrive in the Hybrid Economy* (New York: Penguin, 2009)
- Orly Lobel, 'The Law of the Platform', *Minnesota Law Review* 101 (2016), 87–166
- Viktor Mayer-Schönberger and Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work, and Think* (New York: Eamon Dolan/Houghton Mifflin Harcourt, 2013)
- Fenwick McKelvey and Robert Hunt, 'Discoverability: Toward a Definition of Content Discovery through Platforms', *Social Media + Society*, January/March (2019), 1–15
- Merriam-Webster, 'The Real Story of "Fake News"', available at: <https://www.merriam-webster.com/words-at-play/the-real-story-of-fake-news>
- Persephone Miel and Robert Farris, *News and Information as Digital Media Come of Age* (Cambridge, MA: The Berkman Center for Internet and Society, 2008)

- Evgeni Moyakinea and Alexander Tabachnik, 'Struggling to Strike the Right Balance between Interests at Stake: The "Yarovaya", "Fake news" and "Disrespect" Laws as Examples of Ill-Conceived Legislation in the Age of Modern Technology', *Computer Law and Security Review* 40 (2021), <https://doi.org/10.1016/j.clsr.2020.105512>
- Philip M. Napoli, 'Persistent and Emergent Diversity Policy Concerns in an Evolving Media Environment', in Sean A. Pager and Adam Candeub (eds), *Transnational Culture in the Internet Age* (Cheltenham: Edward Elgar, 2012), 165–181
- Philip M. Napoli, 'On Automation in Media Industries: Integrating Algorithmic Media Production into Media Industries Scholarship', *Media Industries Journal* 1 (2014), 33–38
- Philip M. Napoli, 'Social Media and the Public Interest: Governance of News Platforms in the Realm of Individual and Algorithmic Gatekeepers', *Telecommunications Policy* 39 (2015), 751–760
- Philip M. Napoli et al., *Assessing Local Journalism: News Deserts, Journalism Divides, and the Determinants of the Robustness of Local News* (New Brunswick, NJ: News Measures Research Project, 2018)
- Georg Nolte, 'Hate-Speech, Fake-News, das Netzwerkdurchsetzungsgesetz und Vielfaltsicherung durch Suchmaschinen', *Zeitschrift für Urheber- und Medienrecht* 61 (2017), 552–565
- Dawn C. Nunziato, 'The Marketplace of Ideas Online', *Notre Dame Law Review* 94 (2019), 1519–1583
- Jan Oster, *European and International Media Law* (Cambridge: Cambridge University Press, 2017)
- Eli Pariser, *The Filter Bubble: What the Internet Is Hiding from You* (New York: Viking, 2011)
- Oreste Pollicino, *Judicial Protection of Fundamental Rights on the Internet* (Oxford: Hart, 2021)
- Steven Poole, 'Before Trump: The real History of Fake News', *The Guardian*, 22 November 2019, available at: <https://www.theguardian.com/books/2019/nov/22/factitious-taradiddle-dictionary-real-history-fake-news>
- Julie Posetti and Alice Matthews, *A Short Guide to the History of 'Fake News' and Disinformation* (Washington, DC: International Center for Journalists, 2020)
- Peter Roudik et al., *Initiatives to Counter Fake News in Selected Countries: Argentina, Brazil, Canada, China, Egypt, France, Germany, Israel, Japan, Kenya, Malaysia, Nicaragua, Russia, Sweden, United Kingdom* (Washington, DC: The Law Library of Congress, 2019).
- Matthew Sag, 'Internet Safe Harbors and the Transformation of Copyright Law', *Notre Dame Law Review* 93 (2018), 499–564
- Florian Saurwein et al., 'Governance of Algorithms: Options and Limitations', *info* 17 (2015), 35–49
- Florian Saurwein and Charlotte Spencer-Smith, 'Combating Disinformation on Social Media: Multilevel Governance and Distributed Accountability in Europe', *Digital Journalism* 8 (2020), 820–841
- Olivier De Schutter, *International Human Rights Law*, 2nd edn. (Cambridge: Cambridge University Press, 2014)
- Jacob Soll, 'The Long and Brutal History of Fake News', *POLITICO Magazine*, 18 December 2016, available at: <http://politi.co/2FaV5W9>
- Gerald Spindler, 'Copyright Law and Internet Intermediaries Liability', in Tatiana-Eleni Synodinou et al. (eds), *EU Internet Law in the Digital Era* (Berlin: Springer, 2020), 3–25
- Cass R. Sunstein, *Echo Chambers: Bush v. Gore Impeachment, and Beyond* (Princeton: Princeton University Press, 2001)
- Cass R. Sunstein, *Republic.com 2.0* (Princeton: Princeton University Press, 2007)
- Cass R. Sunstein, *Going to Extremes: How Like Minds Unite and Divide* (Oxford: Oxford University Press, 2009)
- Cass R. Sunstein and Adrian Vermeule, 'Conspiracy Theories: Causes and Cures', *The Journal of Political Philosophy* 17 (2009), 202–227
- Edson C. Tandoc Jr., Zheng Wei Lim and Richard Ling, 'Defining "Fake News"', *Digital Journalism* 6 (2018), 137–153.

- Louis W. Tompros, Richard A. Crudo, Alexis Pfeiffer and Rahel Boghossian, 'The Constitutionality of Criminalizing False Speech Made on Social Networking Sites in A Post-Alvarez, Social Media-Obsessed World', *Harvard Journal of Law and Technology* 31 (2020), 66–109
- Ioanna Tourkochoriti, 'Speech, Privacy and Dignity in France and in the U.S.A.: A Comparative analysis', *Loyola L.A. International and Comparative Law Review* 38 (2016), 101–182
- Heidi Tworek and Paddy Leerssen, 'An Analysis of Germany's NetzDG Law', *Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression* (2019)
- UK House of Commons, Disinformation and 'Fake News', Final Report of the Digital, Culture, Media and Sport Committee, 2019
- 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 Freedom of Expression and 'Fake News', Disinformation and Propaganda, FOM.GAL/3/17, 3 March 2017
- Bart van der Sloot, Dennis Broeders, and Erik Schrijvers (eds), *Exploring the Boundaries of Big Data* (Amsterdam: Amsterdam University Press, 2016)
- Claire Wardle, *Understanding Information Disorder* (London: First Draft, 2019)
- Simon Wechsler, 'The Right to Remember: The European Convention on Human Rights and the Right to Be Forgotten', *Columbia Journal of Law and Social Problems* 49 (2015), 135–165
- Mike Wendling, 'The (Almost) Complete History of "Fake News"', *BBC News*, 22 January 2018, available at : <https://www.bbc.co.uk/news/blogs-trending-42724320>
- Richard S. Whitt, 'A Deference to Protocol: Fashioning a Three-dimensional Public Policy Framework for the Internet Age', *Cardozo Arts and Entertainment Law Journal* 31 (2013), 689–768
- Richard S. Whitt, 'Old School Goes Online: Exploring Fiduciary Obligations of Loyalty and Care in the Digital Platforms Era', *Santa Clara High Technology Law Review* 36 (2019), 75–131
- WHO, Coronavirus disease 2019 (COVID-19), Situation Report 45, 5 March 2020, available at: [https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200305-sitrep-45-covid-19.pdf?sfvrsn=ed2ba78b\\_4](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200305-sitrep-45-covid-19.pdf?sfvrsn=ed2ba78b_4)
- WHO et al., Managing the COVID-19 Infodemic: Promoting Healthy Behaviours and Mitigating the Harm from Misinformation and Disinformation, Joint Statement by WHO, UN, UNICEF, UNDP, UNESCO, UNAIDS, ITU, UN Global Pulse, and IFRC, 23 September 2020, available at: <https://www.who.int/news/item/23-09-2020-managing-the-covid-19-infodemic-promoting-healthy-behaviours-and-mitigating-the-harm-from-misinformation-and-disinformation>
- Patrick Zurth, 'The German NetzDG as Role Model or Cautionary Tale? Implications for the Debate on Social Media Liability', *Fordham Intellectual Property, Media and Entertainment Law Journal* 31 (2021), 1084–1153