Conclusions: Regulatory Responses to Communication Platforms: Models and Limits

Judit Bayer, Bernd Holznagel, Päivi Korpisaari, Lorna Woods (eds)

1. Communication Platforms

In this book, we focused on "communication platforms". Platforms have come to define and dominate several areas of social existence, primarily commerce and communication. Ample research has been published about how social media – in particular combined with the use of a smartphone – changed the communication habits of individuals. The accumulation of these individual actions and habits have resulted measurable changes in societies and politics. We were interested in how platforms effect public communication around the world – as opposed to their market and economic effect – and what the regulatory responses have been in different jurisdictions.

In this, we recognise that there is no one agreed legal definition of "communication platform", and that there may be some variations in the scope of services considered to fall within this category from jurisdiction to jurisdiction. Rather, we proceed on the basis that platforms are services that organise and distribute the information-based content of third parties to a potentially large audience. The value added by platforms is the service of content ranking, personal content recommendations, prioritising and deprioritising, and other currently developing services. The latter activity is what makes platforms so powerful in forming the public discourse. This facilitating action is less than the editorial activity of traditional media service providers, but more than mere "dissemination of information" which does not express the potential of influence and manipulation that is inherent nature of the online platform activity.

At the moment, platforms have some very different definitions, provided by legal instruments that approach different aspects of online platforms, and set different aims. Such are the so-called „platform-to-business regu-

lation"[1] which focused on the commercial angle of online platforms,[2] the draft Digital Services Act,[3] and the draft Digital Markets Act. The German Media State Treaty approaches online platforms from the perspective of public opinion building and defines them ("media intermediary," Medienintermediär) as an online service[4] that aggregates, selects and presents for the general public among others also journalistic-edited content, without combining them into a complete supply.[5]

2. *Effects of modern platform economy on public communication*

In its influential decision from July 21, 2021, the German Federal Constitutional Court analysed the impact of the modern network and platform economy on the process of public opinion-forming as follows:

---

1 Regulation (EU) 2019/1150 of the European Parliament and of the Council of 20 June 2019 on promoting fairness and transparency for business users of online intermediation services (Text with EEA relevance), OJ L 186, 11.7.2019, p. 57.

2 Regulation (EU) 2019/1150 of the European Parliament and of the Council of 20 June 2019 on promoting fairness and transparency for business users of online intermediation services (Platform to business regulation), Recital 1.

3 Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC. Where "dissemination to the public" means making information available, at the request of the recipient of the service who provided the information, to a potentially unlimited number of third parties. Article 2. (h-i). The draft Digital Markets Act uses the expression „platform services" without offering a definition. Instead, it lists the „core platform services" as examples. (Proposal for a Regulation of the European Parliament and of the Council on contestable and fair markets in the digital sector (Digital Markets Act) COM/2020/842 final

4 So-called „telemedium" which is any electronic information and communication service, unless they are telecommunications services under section 3 no. 24 of the Telecommunications Act, consisting entirely of the transmission of signals via telecommunications networks, or telecommunications-based services under section 3 no. 25 of the Telecommunications Act or broadcasting under sentences 1 and 2. See MStV. § 2. 16. and 13.

5 Somewhat confusingly, the German Media State Treaty also uses the word „Medienplattform", however, this term applies to streaming services like Netflix or Amazon Prime: „any telemedium insofar as it combines broadcasting, broadcast-like telemedia or telemedia pursuant to section 19 subsection 1 into an overall offer determined by the provider. The combination of broadcasting, broadcast-like telemedia or telemedia pursuant to section 19(1) is also the combination of software-based applications which essentially serve the direct control of broadcasting, broadcast-like telemedia, telemedia pursuant to section 19(1) or telemedia within the meaning of sentence 1." MStV. § 2. 14.

"[…]Where services are for the most part financed through advertising, they do not necessarily foster journalistic quality; even on the Internet, the large audiences sought by the advertising industry can only be reached by way of programmes that appeal to the masses. In addition, there is the danger that content can be deliberately tailored to users' interests and preferences, also by means of algorithms, which leads to the reinforcement of the same range of opinions. Such services do not aim to reflect diverse opinions; rather, they are tailored to one-sided interests or the rationale of a business model that aims to maximise the time users spend on a website, thus increasing the advertising value of the platform for its clients. […]

This all leads to increased difficulty in the separation of fact from opinion, content from advertisement, as well as to new uncertainties regarding the credibility of sources and assessments. Individual users themselves must now process and assess the information provided by the mass media, which would traditionally have passed through the filter of professional selection in the spirit of responsible journalism. […]."[6] In conclusion, the Court attributes the described changes in the process of forming public opinion – such as the difficulty in the separation of fact from opinion, new uncertainties regarding the credibility of sources and assessments, new burden on individual users to assess the information provided by the internet and social media – to a business model of the platforms, which is financed by advertising and thus has to generate high attention for the content. Maximising attention is achieved through the use of algorithms that address groups of users based on their behaviour and thus inherently carry the risk of manipulation. While for the one-to-many traditional mass media the financing by advertising and the selection of information through gatekeepers (journalists, publishers, broadcasters) were two distinct functions, for the platforms these two functions are governed by the same tool: algorithms are used to optimise the allocation of advertising, and the allocation of content as well. The logic is the same: to generate maximum attention for advertisements. While this logic also existed previously in the traditional mass media, it is realised at a higher efficiency rate with the new characteristics of platform communication. Some of these characteristics are entirely new, like the vanished entry barrier (see point a. below). Others are old features with an enhanced power.

a. *No gatekeeping:* first, contrary to public communication as we knew it in the 20th century, entry barriers vanished with the emergence

6 BVerfG, Judgment of the First Senate of 20. July 2021 - 1 BvR 2756/20 -, Rn. 1-119, http://www.bverfg.de/e/rs20210720_1bvr275620.html.

of platforms. Platforms provide a simple and user-friendly interface which allows anyone to publish content even without literacy (e.g., pictures, videos, sound or simply sharing others' content). All content that is published has the potential to reach a global public. In contrast, content that was meant to reach the public had gone through several layers of filtering in the pre-internet age: owners, editors, journalists kept the public communication under their control. The publishing system naturally enforced a certain financial and educational barrier. Platforms have taken over only some of the gatekeeping roles, the extent of this is still under discussion by policymakers, legislators and platforms themselves.[7]

b. *Personal data:* Second, platforms' activity is driven by personal data. As a primary tool to improve their performance, they collect, aggregate and utilise data, for example in order to optimise their ranking, targeting and recommending systems. It is personal data which drives the placement of advertisements, which is our third point.

c. *Attention-driven advertising:* the competition for the audience's attention has always been the goal of public communication. This has also been a widely criticised pitfall of commercial media. The advertisement-financed content offer's main goal was to maximise the number of financially solvent viewers, which, according to the German Federal Constitutional Court, led to a reduction of content quality.[8] Discussing complex topics would have resulted a loss in audience, therefore priority was given to general themes, and easily accessible content.[9] The goal for platforms is the same, but the means to the end, and the consecutive result are different. Polarising themes can be targeted at susceptible audiences. In absence of the entry barriers (see point a.), this becomes a race to the bottom. With the help of algorithms (see in d.) finding the right person for the right content can be perfected, and thereby the attention of users can be exploited in a much more effective way than by traditional commercial media. The format of some platforms leads to shorter communications, which may also be less sophisticated in analysis. Some communication tools, e.g., emojis

7 This gatekeeping role in public communication is not to be confused with the emerging gatekeeping role of platforms in regard of platform communication.

8 BVerfG, Judgment of the First Senate of 20. July 2021 - 1 BvR 2756/20 -, Rn. 1-119, http://www.bverfg.de/e/rs20210720_1bvr275620.html.

9 See among others, for example: McChesney, Robert W. "Corporate Media and the Threat to Democracy", *Penguin Random House*, 1997; Curran, James, "Media and Democracy", *Routledge*, 2011.

and 'likes', can lead to swift communication, but may lead to many different interpretations and the risk of misunderstanding. A further concern is that the constant strive for positive feedback (likes, upvotes and other signs of public approval) affects the types of content produced; research suggests that users are more likely to share sensational disinformation than truthful content.[10] With the ubiquitous presence of social media through our smartphones, this brings about the problem of information overconsumption. Attention is becoming a scarce resource and not all users are capable to manage it wisely.

d. *Algorithms and AI:* the governance of content distribution, personal data aggregation, advertisement auctions, targeting, recommending, ranking and many more actions on which social media is built, would not be possible without algorithms and AI solutions. Automation is also applied in content moderation, although human supervision appears still inevitable in that regard. AI has also appeared as „artificial users", social bots, ad bots, pol (political) bots and trading bots,[11] which are potential influencers of public communication trends. In this sense AI has the potential to manipulate public opinion building.[12]
e. *Concentration:* finally, the public communication sphere is dominated by some giant companies. There are thousands of small companies, but a small number of large companies hold the biggest market shares. The reason for this development are network effects of communication platforms.[13] This phenomenon has been deeply analysed in the

10 Vosoughi, Soroush, Roy, Deb, Aral, Sinan. „The spread of true and false news online." *Science* 09 Mar 2018. Vol. 359, Issue 6380, pp. 1146-1151. DOI: 10.1126/science.aap9559. See also: Islam, A., Laato, S., Talukder, S., & Sutinen, E. (2020). Misinformation sharing and social media fatigue during COVID-19: An affordance and cognitive load perspective. Technological forecasting and social change, 159, 120201. https://doi.org/10.1016/j.techfore.2020.120201.

11 Caprolu, Maruantoni, Cresci, Stefano, Raponi, Simone, Di Pietro, Roberto. „New Dimensions of Information Warfare: The Economic Pillar—Fintech and Cryptocurrencies." *Risks and Security of Internet and Systems: 15th International Conference, CRiSIS 2020, Paris, France, November 4-6, 2020, Revised Selected Papers*. Springer Nature, 2021. p. 3.

12 BVerfG, Judgment of the First Senate of 20. July 2021 - 1 BvR 2756/20 -, Rn. 1-119, http://www.bverfg.de/e/rs20210720_1bvr275620.html.

13 Recital 55 Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC. See also: Gillespie, Tarleton. "Content Moderation, AI, and the Question of Scale." Big Data & Society, (July 2020). https://doi.org/10.1177/2053951720943234. Ofcom, "Use of AI in online content moderation". 2019 Report. https://www.ofcom.org.uk/__data/assets/pdf_file/0028/157249/

telecommunications sector.[14] The platform giants have unprecedented numbers of users (e.g., 2,85 billion for Facebook in March 2021[15]) and impressive profit rates (30-40% 2020-2021[16]). On the one hand, this audience reach is theoretic: typically, not all users see the same content. How many people see a certain piece of content is defined by many factors. However, with the help of algorithms and the available personal data, the platforms are in the position to influence this reach. On the other hand, the power of the giant companies is unprecedented in public communication – in comparison with traditional media companies –, with substantial consequences on their lobbying power against regulatory initiatives. Besides, these few giant platforms increasingly act as gatekeepers between business users and end users, and the misuse of their dominant position can be suspected.[17] The significant difference in power between small and large platforms justifies the differentiated treatment of large platforms, as it is envisaged in the draft Digital Services Act.

3. *Platform harms*

The impact that platforms exercise on public communication, cannot be easily categorised. All induced changes carry elements that can be evaluated positively or negatively. The circumstances and the context define whether a certain way of usage causes positive or negative effects for a certain individual, or a group of people.[18] For example, the spread of conspiracy theories in the context of the Covid-19 pandemic is celebrated

---

cambridge-consultants-ai-content-moderation.pdf. See further: Bradshaw, S. (2019). Disinformation optimised: gaming search engine algorithms to amplify junk news. *Internet Policy Review*, 8(4). http://dx.doi.org/10.14763/2019.4.1442.

14 Kühling, Jürgen, Schall, Tobias, Biendl, Michael, "Netzwerkeffekte ausführlich dargestellt", Telekommunikationsrecht, no. 2, 2014, Pages 50-53.

15 Statista, *Number of monthly active Facebook users* worldwide, https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/

16 Facebook Profit Margin (Quarterly): 36,29% for March 31, 2021. YCharts. https://ycharts.com/companies/FB/profit_margin.

17 European Commission, "Antitrust: Commission opens investigation into possible anticompetitive conduct of Facebook". 4 June 2021. https://ec.europa.eu/commission/presscorner/detail/en/ip_21_2848.

18 For example, the possibility of anonymous content sharing and getting connected to like-minded people brought up the me-too movement, and helped victims of stigmatising crimes to speak and find support. The same features which help

by some as an expression of their freedom of expression, while others see it as a danger to themselves and to public health. When a piece of content is clearly criminal, like child pornography or terroristic content, there is often a broad consensus in society and across legal instruments that it should be removed. However, it is not always easy to evaluate, whether content is criminal, illegal under another law, or legal. For example, defamation is criminal in some states and a civil wrong in others; beyond that, it is contextual and its evaluation might depend on several defences.[19] Hate speech can be used to cover a vast swathe of comment – from mere slurs at one end of the scale to incitement to genocide at the other.[20] Within this range, the placement of the boundary for criminal offences may lie at different points in different states. Moreover, it is context-dependent in most cases.

The example of the draft Digital Services Act illustrates some of the difficulties in this area. Article 2(g) of the draft Digital Services Act speaks about 'illegal content' meaning "any information, which, in itself or by its reference to an activity, including the sale of products or provision of services is not in compliance with Union law or the law of a Member State, irrespective of the precise subject matter or nature of that law". Recital 12 DSA further clarifies that the term 'illegal' is a broad one. It may refer to information, that under the applicable law is either itself illegal, or which relates to activities that are illegal.[21] For a digital service provider, it could be difficult to judge, whether they should block or delete information from a platform due to "illegality" – especially because the "illegality" might also differ in content from one Member State to another. The removal raises technical and procedural questions that impact users' rights, such as notification of the content provider, and the possibility to put back the content.[22] Ultimately a court should decide

---

marginal groups to organise themselves also foster political extremism, hate speech or hate crime.

19 Law Commission of Ontario, *Defamation Law in the Internet Age*, March 2020 at 77, <www.lco-cdo.org/wp-content/uploads/2020/03/Defamation-Final-Report-Eng-FINAL-1.pdf> (accessed 15 July 2021).

20 United Nations Strategy and Plan of Action on Hate Speech, May 2019 at 12, https://www.un.org/en/genocideprevention/hate-speech-strategy.shtml (accessed 13 August 2021).

21 See Rec 12 of the draft DSA that makes clear that the term ‚illegal' is a broad one.

22 See also *Glawischnig-Piesczek v Facebook Ireland Limited* (C-18/18), where the CJEU judged that EU law does not preclude a host provider from being ordered to remove identical and, in certain circumstances, equivalent comments previously declared to be illegal.

about illegality, while supervision authorities and agencies may play a role in notifying relevant illegal content and ordering their removal.[23] Jurisdictional differences complicate the picture for the global platforms.

The position as regards other than criminal content is even more complex. Some information could be termed 'illegal' but not criminal, because (depending on jurisdiction) they are contrary to other types of law: e.g., misleading advertising. And, a wide range of information can be termed 'harmful', i.e., content that does not trigger a legal response outside the platform environment. This last category of content might still be dealt with by platforms enforcing their community standards. Increasingly, however, there are concerns about content (e.g., COVID denial) that in offline context is potentially harmful but has little opportunity to spread, however, in a platform environment is accessible to a large audience and in many cases actively promoted by the platform systems. The problem occurs from the interplay between the content and the platform's distribution system (and their features (a)-(e) noted above). This has led to suggestions that the "online ecosystem" as such should be regulated (see more on this below). Where human rights are in issue – as here – it must be remembered that state measures must always be specific and proportionate. Legal measures may only exist if there are legitimate reasons (such as the protection of minors, fairness in business transactions, protection of reputation). It is therefore advisable to determine precisely which online harms require which countermeasures for which reasons.

It is also important to take into consideration, that according to the practice of the ECtHR, freedom of expression also applies to expressions that offend, shock or disturb, including untrue facts. Therefore, all limitations to freedom of expression have to be construed strictly, and the need for any restriction must be established convincingly.[24] However, states have positive obligations to ensure protection of privacy, and also the chances of a plural information environment. According to the case law of the ECtHR, it may be justified to restrict expression for these purposes, e.g., ECtHR upheld restrictions against misleading advertisements,[25] and

23 See also: Advocate General's Opinion in Case C-401/19, Poland v Parliament and Council. https://curia.europa.eu/jcms/upload/docs/application/pdf/2021-07/cp210138en.pdf.

24 For example, *Hertel v Switzerland* App no 25181/94 (ECtHR, 25 August 1998) Reports of Judgments and Decisions 1998-VI; *Steel and Morris v the United Kingdom* App no 68416/01 (ECtHR, 15 December 2005) ECHR 2005-II; *Stoll v Switzerland* [GC] App no 69698/01 (ECtHR, 10 December 2007) ECHR 2007-V;.

25 Hertel v Switzerland (2002) App. No. 53440/99, 17 January 2002.

against a campaign by the Raëlien Movement which fostered believes that life on Earth was created by extraterrestrials, among others.[26]

The high volume of hate speech and disinformation is currently seen as shaking the foundations of our democracies. Political disinformation has been seen capable to influence elections, induce riots, lynching, mobbing, and even genocide.[27] Health disinformation may cost lives and hamper the defence against deadly diseases.[28] Hate speech and harassment against vulnerable groups intimidate their victims and have induced violent attacks against several of them.[29] Beyond the actual harms in the life and safety of the victims, the mentioned content has been causing fissures in the social cohesion and the functioning of democracy.[30] Truth and trust have become concepts that we have become unable to authentically identify.

The low entry barrier into public communication enabled by social media opened the possibility for masses of people to let their voice heard, and react to others. In a quest for popularity, some leaders target people from whom they hope the widest support. Vulnerable minorities are easy target points for both as scapegoats (inciting other users to attack them) and as targets of misleading advertising. Lacking editorial responsibility, and

---

26 *Mouvement Raëlien Suisse* v *Switzerland* (2013) 56 EHRR 14 para 62.

27 116th Congress Senate Report of the Select Committee on Intelligence US Senate on Russian Active Measures Campaigns and Interference in the 2016 US Election. Volume 2. Russia's Use of Social Media With Additional Views. https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume2.pdf.; House of Commons, Digital, Culture, Media and Sport Committee. Disinformation and 'fake news': Final Report. 18 February 2019. https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/1791.pdf.; UN Human Rights Council Report of the independent international fact-finding mission on Myanmar. A/HRC/39/64. https://www.ohchr.org/Documents/HRBodies/HRCouncil/FFM-Myanmar/A_HRC_39_64.pdf.

28 Bayer, Judit, Holznagel, Bernd, Lubianiec, Katarzyna,, Pintea, Adela,, Schmitt, Josephine B, Szakács, Judit, Uszkiewicz, Erik, (2021) Disinformation and propaganda: impact on the functioning of the rule of law and democratic processes in the EU and its Member States - 2021 update. EP/EXPO/INGE/FWC/2019-1/LOT6/R/07.

29 Bayer, Judit and Bárd, Petra: Hate speech and hate crime in the EU and the evaluation of online content regulation approaches. Study for the European Parliament, Policy Department C: Citizens' Rights And Constitutional Affairs. 2020. ISBN 978-92-846-6902-8 | doi:10.2861/28047.

30 McKay, Spencer and Chris Tenove. "Disinformation as a Threat to Deliberative Democracy." Political Research Quarterly, (July 2020). https://doi.org/10.1177/1065912920938143. ; See also: Luttrell, Regina - Xiao, Lu – Glass, Jon (eds), "Democracy in the Disinformation Age. Influence and Activism in American Politics." Routledge. 2021.

pillars of truth, spreading of populistic disinformation becomes easy. Technological capacity allows the amplification and manipulation of messages, e.g., through the use of bots, trolls, disinformation networks or deep fakes. The attention-based advertising model advantages sensational content and disadvantages rational presentation of facts. In our days, this mechanism defines all public communication, including political communication.

The new social gap appears along the lines of rational thinkers and believers. On the one hand, it is important that all citizens feel represented in democracies, and all people have the right to believe and think what they do. On the other hand, social functioning cannot be based on false facts and conspiracy theories. All the freedom of expression theories have been based on the presumption that people are rational human beings and that in an open discussion, truth will prevail.[31] A minority of extremists can and should be tolerated by the majority, and their contest of ideas is supposed to lead to better solutions. However, if more than a small minority follows extremist ideas, that is bound to disrupt the functioning of democracy. Thus, the challenge of our age is to turn the tide: to reduce the spreading of false beliefs and conspiracy theories without prohibiting them and without stigmatising the people who believe in them. The goal should be to reduce their representation to a level which is tolerable in a constitutional democracy.

Therefore, the action ground ought to be the distribution logic of this platform-based public communication system, rather than fighting against certain content or the people who like and share them. One way could be that certain rules and conditions were amended so that verified information has better chances to be accessed than disinformation. But which rules and conditions would those be, and how would the truthfulness of information be verified in a rapid communication environment?

The possibilities offered by the rapidly developing platform technology are complex for legislative policy making. Legislation takes years to get finalised, and the development rushes by. Freedom of pursuing business, and other freedoms are also factors to be respected. Against this background, there are strong forces in Europe to develop a counterweight, safeguarding diversity and providing guidance in the post-truth information environment. As the German Constitutional Court argues, public broadcasters are even becoming more significant in "times of increased complex information on the one hand and one-sided representations, fil-

31 Mill, John Stuart: On Liberty. Boston. 1863. p. 50-58.

ter bubbles, fake news, deep fakes on the other".[32] The British regulator Ofcom, similarly, called for updating the system of public broadcasting.[33] An in-depth consultation has been pursued exploring the possibilities of how to adapt the system to the changing informational environment.[34] Meanwhile, the Finnish government proposed a bill, limiting the Finnish Broadcasting Company (Yle) to publish longer texts only in support of video or audio broadcast, rather than independently.[35] The move is to preserve fair competition between commercial media and Yle. According to the director of Yle, the change can also foster reform and strengthening of Yle.[36] Self-regulation appeared to be a route that builds on the know-how of those who best understand what platforms are able to do: platforms themselves. Platforms did indeed large efforts to introduce measurements in their communication systems to reduce the visibility of disinformation and hate speech.[37] The assessment of the self-regulation showed that the efforts were partly successful, but they were diverse across platforms and countries and also incalculable. The European Regulators' for Audio-visual Media Services emphasised the inconsistent application and the insufficiency of the oversight mechanism.[38]

---

32 BVerfG, Judgment of the First Senate of 20. July 2021 - 1 BvR 2756/20 -, Rn. 1-119, http://www.bverfg.de/e/rs20210720_1bvr275620.html. For further reference, see: report of the Enquête Commission on Artificial Intelligence of the German Parliament (Bundestag) of 28 October 2020, BTDrucks 19/23700, p. 447 ff.).

33 Ofcom „Ofcom calls for stronger system of public service media fit for the digital age." July 15 2021. https://www.ofcom.org.uk/about-ofcom/latest/media/media-releases/2021/stronger-public-service-media-system-for-digital-age

34 Ofcom „Small Screen, Big Debate. Consultation. The Future of Public Service Media." December 8. 2020. https://www.smallscreenbigdebate.co.uk/__data/assets/pdf_file/0032/208769/consultation-future-of-public-service-media.pdf

35 Yle „Gov't aims to limit Yle web publications." June 16. 2020. https://yle.fi/uutiset/osasto/news/govt_aims_to_limit_yle_web_publications/11405119

36 Ibid.

37 Washington Post "Facebook says it has taken down 7 million posts for spreading coronavirus misinformation. The company also labeled 98 million posts with warning notices about coronavirus misinformation between April and June." August 11. 2020. https://www.washingtonpost.com/technology/2020/08/11/facebook-covid-misinformation-takedowns/. More recently the Centre for Countering Digital Hate's report, Failure to Protect, https://www.counterhate.com/failuretoprotect, suggests that 84% of antisemetic posts were not taken down.

38 European Commission Staff Working Document, 'Assessment of the Code of Practice on Disinformation - Achievements and areas for further improvement', SWD(2020) 180 final, p. 22.

Under the current scheme, platforms have a substantial income from the spreading and flourishing of disinformation and extreme content. A systematic restructuring of the communication patterns would result that platforms lose part of their revenues, unless they also restructure their income base. Considering the stellar profits that giant platforms make, obviously there is ample room for manoeuvre in this area. But expecting that platform companies would proactively cut their own profit appears reasonable only if they are given clear expectations with the possibility of enforcement. The worldwide attempts that are reflected in this book, to draft some kind of control on social media platforms, can be interpreted as a signal that the time is ripe for this move.

4. How to deal with the harms?

This book highlights several snapshots of legal approaches and instruments which aim at dealing with the dangers caused by online platforms. Some states are dealing with issues through the lens of data protection (e.g. Russia), or by focussing on market regulation (to some extent, the USA). Both these in some way relate to the business model of the platforms, as do proposals that focus on the design of the platforms. There are also approaches that focus on content regulation. Below, we attempt to typify and order the approaches that we have encountered during the project.[39]

a. Defining a general duty of care standard

We know now that the inherent structure of platform communication carries the risk of distorting the social discourse. The individual violations of rights cannot reflect accurately the systemic distortion of the communication scene. Addressing only the individual violations of law, or even the individual pieces of harmful content, will not change the systemic harms. While these systemic harms may be indirect, their effects are more than

39 The workshop series on Hate speech and platform regulation included seven workshops and several speakers who were not included in the volume. Report on the project in Bayer, Judit, Kalbhenn, Jan, „Masse und Macht – Auf der Suche nach Regeln für digitale Kommunikationsplattformen“, ZUM – Zeitschrift für Urheber- und Medienrecht, No. 4 (2021).

subtle[40] and threaten the operation of democracies, the basic foundation of which is free, but also rational, discourse on common matters. Therefore, there is good reason to view platforms, especially social media platforms as systems which carry an inherent systemic risk, like rail, automobile, or powerplants. Their operators should be aware of the risks and do all necessary efforts in their competence to minimise those risk.

This approach is followed by the United Kingdom's „Duty of Care" principle found in the draft Online Safety Bill, and the EU's draft Digital Services Act's „risk assessment" obligation, with some meaningful differences. First, as opposed to the scheme in relation to rail and automobiles, platforms' liability for the damage caused through the platform as a vehicle, is generally exempted. They do not bear direct liability to cover the losses caused by illegal content, for example. But they still might be made *responsible* by law, to design and apply the preventive measures to minimise the risk.[41] The UK model includes an obligation to *design a system* that allows for content notified as illegal to be taken down swiftly.[42] As regards children, platforms are under an obligation to mitigate and manage the risks of harm and in some instances, using system design, prevent children from coming into contact with specified types of content.[43] Although the UK has left the EU, it currently maintains the immunity provisions derived from the e-Commerce Directive. The interplay between the two sets of provisions is not yet known. Whereas, the EU's draft Digital Services Act strictly orders removal of illegal content when platforms are notified of them, as a condition of their exemption from liability. The "due diligence" obligations apply to all other issues, including procedures, transparency, dealing with harmful content, and more.

40 Ibbetson, Connor "Where do people believe in conspiracy theories?" *YouGov Cambridge Globalism Project*. 18 January, 2021. https://yougov.co.uk/topics/international/articles-reports/2021/01/18/global-where-believe-conspiracy-theories-true.

41 On delineation of liability and responsibility see: Chapter 5.1. by Sarah Hartmann in: Bayer, Judit, Katsirea, Irini, Batura, Olga, Holznagel, Bernd, Hartmann, Sarah, Lubianiec, Katarzyna. The fight against disinformation and the right to freedom of expression. (Brussels: European Parliament, 2021) p. 59-63. https://www.europarl.europa.eu/RegData/etudes/STUD/2021/695445/IPOL_STU(2021)695445_EN.pdf. The draft Online Safety Bill imposes an obligation to operate a system that minimises the presence of illegal content and to mitigate against the likelihood of children encountering content of a type assessed to be harmful to them.

42 Clause (3)(d) draft Online Safety Bill.

43 Clause 10(2) and (3) draft Online Safety Bill.

b. Duty of standards in specific areas (sectors)

In the broader picture, it is important to note that the topics that need to be assessed as systemic risks by the platform providers, are manifold. Some of these topics are also regulated by separate acts, such as the use of artificial intelligence, data protection, the protection of children and advertising. Others may be regulated by some states, but are often left to self-regulation, such as hate speech and disinformation. However, recently, many countries took countermeasures in this area. For example, the Canadian government has left behind its reservations to regulate internet communication and has introduced a comprehensive law against hate speech. Intensive discussions are also taking place in other countries, on how effective action can be taken in particular against hate speech and disinformation, as it was reflected in our workshops (in Japan, Singapore, India, etc.).

A further systemic risk which is not left to self-assessment and self-regulation, but falls entirely in the realm of state regulation, is market concentration of platform operators and the risk to the fair economic competition.[44] In the USA in particular, there is intensive discussion about whether limiting the economic power of the large platforms could be an important prerequisite not only for more competition, but also for effectively combating hate communication and disinformation. There is also intensive discussion here about whether interoperability obligations, as we know them from telecommunications law, can contribute to increasing the number of communication platforms such as Facebook.[45] There is a more general concern about the operation of competition law with regard to the super-dominant tech companies. The Digital Markets Act envisages special rules to be applied to digital gatekeepers; the UK is considering similar measures. Germany[46] has already passed a new law. The common

44 The draft Digital Services Act approaches the areas of risk assessment from the perspective of activities: besides (a) dissemination of illegal information, it relates to (b) any negative effects for the exercise of the fundamental rights to respect for private and family life, freedom of expression and information, the prohibition of discrimination and the rights of the child; and (c) the intentional manipulation and exploitation of their service, with an actual or foreseeable negative effect on the protection of public health, minors, civic discourse, or actual or foreseeable effects related to electoral processes and public security.

45 See more in Chapter 1.4. Policy Developments in the USA to Address Platform Information Disorders by Sarah Hartmann.

46 Section 19a Kartellgesetz 2020. (German Competition Act).

theme is that there is some element of ex ante obligations, an approach which may have been borrowed from the telecommunications regime.

c. *Enforcing the duty of care standard: self-regulation, co-regulation or state supervision*

This leads us to the second question: provided that the industry would design the measures to mitigate the harms, and given that the actors are not directly liable for direct harms, how is the efficacy of the system supervised and enforced? The direct harms are not the liability of platforms, and the indirect harms are not measurable – how to ensure that the risk minimising measures are successful? In the United Kingdom, this task would be allocated to Ofcom, the converged authority for media, telecommunication and post.[47] As a regulator, it would be competent to supervise and impose orders on the platform operators. The European Union's draft Digital Services Act provides for a complex set of supervisory and compliance measures. National competent authorities, with designated Digital Services Coordinators would have wide powers to investigate, seek information and impose orders, as well as penalties. In case the national procedure is insufficient (cases defined precisely in the Act), the European Commission may exercise delegated powers to investigate and enforce the Regulation and the relating decisions. Critiques find that authorities' role is exaggerated, in view of the freedom of expression standards which require access to courts. Yet, the Council of Europe has also recommended the establishment of regulatory authorities in the context of the broadcasting sector, albeit with strong emphasis on the independence of such authorities.[48] Provided that ultimate judicial review of the authority's decision remains possible, authorities may need to play a proportionate role in justified restrictions of freedom of expression rights also in the platform environment, considering the abundance of content which would otherwise overload the judicial system. The draft Digital Services Act provides for a set of "due diligence" obligations, which aim at different public policy objectives such as the safety and trust of the recipients of the service, including minors and vulnerable users, protecting the relevant

47 See Chapter 1.3. by Lorna Woods in this book.

48 Recommendation Rec(2000)23 of the Committee of Ministers to member states on the independence and functions of regulatory authorities for the broadcasting sector.

fundamental rights enshrined in the Charter, empowering recipients and other parties.[49] Some of the rules under Chapter III. which sets out the due diligence obligations, are formulated strictly, such as the notice and action procedure (dealing with illegal content) and the transparency obligations.[50] Also, several of the obligations that apply to very large online platforms (Chapter III. Section 4) are straightforward and easily controllable, such as the transparency of recommender systems, additional online advertising transparency, data access, appointing compliance officers, and the further transparency obligations.[51] However, in the case of other obligations, checking adequate compliance may be a complicated endeavour. Such is the obligation to identify, analyse, assess the risks; to put in place reasonable, proportionate and effective mitigation measures; and to have an independent audit.[52] The exact content of these expectations from very large online platforms is left open, to be developed by the industry actors themselves, in particular in the Code of Conduct, and in the Advertising Code of Practice. No enforcement measures are planned in relation to the envisaged code of practice (and the advertising code). There is no clear provision on whether the Coordinator can decide if the measures taken to mitigate the risks are insufficient. This is supposed to be established by the independent audit. However, a negative audit report entails nothing more than the obligation to justify the reasons for not implementing the operational recommendations – and setting out „any alternative measures they may have taken to address any instances of non-compliance identified". At this stage, it is unclear whether the Digital Services Coordinators would have the power to declare that the operator did not adequately justify the reasons for not implementing the recommendations.

Digital Services Coordinators may start their procedure only in case of an infringement of the rules of the Regulation. In that case, they may adopt a decision on the infringement, and request the platform to draw up an action plan. The Digital Services Coordinator may then decide whether the action plan is appropriate, and it may request the platform to subject itself to an additional independent audit, with an appointed auditor.

In sum, the enforcement system of the draft Digital Services Act is very carefully designed, and sets out considerable fines in case of violation of the Act, but the Act defines only the basic obligations of platforms,

49 Recital 34-35. Digital Services Act.
50 Article 10-24. Digital Services Act.
51 Articles 29-33.
52 See Section 4 of Chapter III. See also Section 50. (1).

whereas many details are referred to self-regulation. The self-regulatory codes are passed under the supervision of the European Commission, but the consequences of non-compliance with the Codes are not clarified in the Act.

d. Supervision: allocating competences between competent authorities

When it comes to the supervision and regulation of these – often overlapping – areas of systemic risks, the question of allocating competences between competent authorities emerges. It can already be observed that in particular the data protection authorities, the cartel authorities and the media and telecommunications regulators are arguing about who should be responsible for combating online harms. In the EU, in addition to this problem of horizontal distribution of supervisory responsibilities, there is also the problem of vertical distribution of competences between the European Commission or EU agencies, and the national authorities.

In the UK, Ofcom has responsibility for the range of communications industries, including now video sharing platforms and, when the draft Online Safety Bill comes into force, other social media platforms. It nonetheless needs to work with other regulators – notably the Competition and Markets Authority (dealing with competition and consumer protection), the Information Commissioner's Office (responsible for data protection and freedom of information) and even the Financial Conduct Authority (the financial services regulator). To do this, the Digital Regulators Cooperation Forum has been established. It remains to be seen how effective it will be. Extensive discussions within each jurisdiction are likely to be needed to develop an effective supervisory model.

5. Final remarks: do we need a global regulation?

Many of our expert authors have expressed the view that national regulation is not expected to be successful against the actions of global online platforms. It has even been noted that actual notices and requests by state authorities have been seen to be ignored by giant companies. This leads us to ask whether transnational regulation or international rules would deal with the mentioned social, individual and economic problems more efficiently. However, this has some obstacles. First, national legal frameworks are different, especially when it comes to content regulation.

The draft Digital Services Act plans to overcome this difficulty with the transnational hub of the Digital Services Board and the Commission – a scheme that has been applied in the General Data Protection Regulation before. But still, the evaluation on what is "illegal" and what is permitted, would be defined through national regulations. It may, however, be that regulation of the distribution of content (systems or ecosystem regulation) is marginally less contentious than direct content regulation. Second, the globe is divided in major attitudes towards regulation. China or Russia have vastly different standards than the United States, with Europe and other continents being also divergent. Inspite of these hindrances, there is some hope to come to common denominators provided there is an intention to do so. There are some soft law initiatives being developed at the international level (e.g., OSCE Guidance on AI in content moderation). A regional cooperation between democratic states would be possible and also desirable. Currently, all states appear to keep their eyes on other states, watching what those are initiating to tackle the problems which press so many societies worldwide. Therefore, there is considerable responsibility on the European Union and those states which lay the groundwork for a new regulatory regime.

Bibliography

116th Congress Senate Report of the Select Committee on Intelligence US Senate on Russian Active Measures Campaigns and Interference in the 2016 US Election. Volume 2. *Russia's Use of Social Media With Additional Views*. Washington D.C.: 116th Congress Senate. (2019). https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume2.pdf.

Bayer, Judit, Bárd, Petra. *Hate speech and hate crime in the EU and the evaluation of online content regulation approaches*. Brussels: European Parliament, Policy Department C: Citizens' Rights And Constitutional Affairs, 2020. doi:10.2861/28047.

Bayer, Judit, Holznagel, Bernd, Lubianiec, Katarzyna,, Pintea, Adela, Schmitt, Josephine B, Szakács, Judit, Uszkiewicz, Erik. *Disinformation and propaganda: impact on the functioning of the rule of law and democratic processes in the EU and its Member States - 2021 update*. Brussels: European Parliament, 2021. https://www.europarl.europa.eu/RegData/etudes/STUD/2021/653633/EXPO_STU(2021)653633_EN.pdf.

Bayer, Judit, Kalbhenn, Jan, *Masse und Macht – Auf der Suche nach Regeln für digitale Kommunikationsplattformen*. ZUM – Zeitschrift für Urheber- und Medienrecht, No. 3 (2021): 185-194.

Bayer, Judit, Katsirea, Irin, Batura, Olga, Holznagel, Bernd, Hartmann, Sarah, Lubianiec, Katarzyna. *The fight against disinformation and the right to freedom of expression*. Brussels: European Parliament, 2021. https://www.europarl.europa.eu/RegData/etudes/STUD/2021/695445/IPOL_STU(2021)695445_EN.pdf.

Bradshaw, S. "Disinformation optimised: gaming search engine algorithms to amplify junk news." *Internet Policy Review,* 8(4) (2019). http://dx.doi.org/10.14763/2019.4.1442.

Caprolu, M., Cresci, S., Raponi, S., Di Pietro, R. "New Dimensions of Information Warfare: The Economic Pillar—Fintech and Cryptocurrencies." *Risks and Security of Internet and Systems: 15th International Conference, CRiSIS 2020, Paris, France, November 4-6, 2020, Revised Selected Papers*. Springer Nature (2021): 3.

Centre for Countering. "Digital Hate's report, Failure to Protect." Accessed August 12, 2021. https://www.counterhate.com/failuretoprotect.

Curran, James. *Media and Democracy*. Routledge. (2011).

European Commission. *Antitrust: Commission opens investigation into possible anti-competitive conduct of Facebook*. Brussels: European Commission, 4 June 2021. https://ec.europa.eu/commission/presscorner/detail/en/ip_21_2848.

European Commission. *Assessment of the Code of Practice on Disinformation – Achievements and areas for further improvement*. Brussels: European Commission, 2020. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=69212.

Gillespie, T. "Content Moderation, AI, and the Question of Scale." *Big Data & Society* (July 2020). https://doi.org/10.1177/2053951720943234.

House of Commons, Digital, Culture, Media and Sport Committee. *Disinformation and 'fake news': Final Report.* London: Hose of Commons, Digital Culture, Media and Sport Committee, 2019. https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/1791.pdf.

Ibbetson, C. "Where do people believe in conspiracy theories?" *YouGov Cambridge Globalism Project.* (January 2021). https://yougov.co.uk/topics/international/articles-reports/2021/01/18/global-where-believe-conspiracy-theories-true.

Islam, A., Laato, S., Talukder, S., & Sutinen, E. "Misinformation sharing and social media fatigue during COVID-19: An affordance and cognitive load perspective". *Technological forecasting and social change*, 159, 120201. (2020). https://doi.org/10.1016/j.techfore.2020.120201.

Kühling, Jürgen, Schall, Tobias and Biendl, Michael. "Netzwerkeffekte ausführlich dargestellt" *Telekommunikationsrecht* no. 2 (2014): 50-53.

Law Commission of Ontario, *Defamation Law in the Internet Age*, Ontario: Law Commission, 2020. www.lco-cdo.org/wp-content/uploads/2020/03/Defamation-Final-Report-Eng-FINAL-1.pdf.

Luttrell, Regina, Xiao, Lu, Glass, Jon (eds). "Democracy in the Disinformation Age. Influence and Activism in American Politics." *Routledge.* (2021).

McChesney, Robert W. *Corporate Media and the Threat to Democracy*. Penguin Random House, 1997.

McKay, S., Tenove, C. "Disinformation as a Threat to Deliberative Democracy." *Political Research Quarterly*, (July 2020). https://doi.org/10.1177/106591292093814 3.

Mill, John Stuart. *On Liberty*. Boston: Tricknor and Fields. (1863).

Ofcom, "Use of AI in online content moderation". (2019). https://www.ofcom.org.uk/__data/assets/pdf_file/0028/157249/cambridge-consultants-ai-content-moderation.pdf.

Statista. "Number of monthly active Facebook users worldwide." Accessed July 29, 2021. https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/.

UN Human Rights Council. *UN Human Rights Council Report of the independent international fact-finding mission on Myanmar*. Geneva: UN Human Rights Council, 2018. https://www.ohchr.org/Documents/HRBodies/HRCouncil/FFM-Myanmar/A_HRC_39_64.pdf.

Vosoughi, S., Roy, D., Aral, S. "The spread of true and false news online." *Science* , Vol. 359, Issue 6380 (March 2018): DOI: 10.1126/science.aap9559.

Washington Post "Facebook says it has taken down 7 million posts for spreading coronavirus misinformation. The company also labeled 98 million posts with warning notices about coronavirus misinformation between April and June." *Washington Post*, August 11, 2020. https://www.washingtonpost.com/technology/2020/08/11/facebook-covid-misinformation-takedowns/.

YCharts. "Facebook Profit Margin (Quarterly): 36,29% for March 31, 2021.". Accessed July 21, 2021. https://ycharts.com/companies/FB/profit_margin.