Mercer Law Review

Volume 73
Number 2 *Lead Articles Edition*

Article 6

3-2022

The Protection of Freedom of Expression from Social Media Platforms

András Koltay

The Protection of Freedom of Expression *from* Social Media Platforms

András Koltay*

I. INTRODUCTION

Social media platforms have overturned the previously known system of public communication. As predicted at the outset, the spread of the public Internet that started three decades ago has resulted in a paradigm shift in this field. Now, anyone can publish their opinion outside the legacy media, at no significant cost, and can become known and be discussed by others. Due to the technological characteristics of the Internet, it might also be expected that this kind of mass expression, with such an abundance of content, would necessitate the emergence of gatekeepers, similar in function to the ones that existed earlier for conventional media. The newsagent, post office, and cable or satellite services have been replaced by the Internet service provider, the server (host) provider and the like. However, no one could have foreseen that the new gatekeepers of online communication would not only be neutral transmitters or repositories but also active shapers of the communication process, deciding on which user content on the Internet they deemed undesirable and deciding which content, out of all the theoretically accessible content, is actually displayed to individual users. Content filtering, deleting, blocking, suspending, and ranking are

*Professor of Law, University of Public Service (Budapest) and Pázmány Péter Catholic University (Budapest). I would like to thank the valuable comments, suggestions and support of RonNell Andersen Jones, Jon Garon, Margaret Hu, Seth Oranburg, Sue Painter-Thorne, Eric J. Segall, Gary Simson, and Russell L. Weaver, participants of the Symposium of the Mercer Law Review on October 8, 2021. Thanks are also due to the organizers and participants of the Media Law & Policy Scholars Conference 2022 (organized at the University of Texas at Austin), especially to the generous discussant of my paper, Professor Rob Kahn. I am also grateful to Édua Reményi, who did a sterling job editing my earlier manuscripts, and to Szabolcs Nagy and to Steven Patrick who helped me with the English adaptation of my original drafts.

all types of active interference with the exercise of users' freedom of speech and practices which also affect the interests of other users in obtaining information. All this became an even greater and more difficult-to-manage issue when, in certain sub-markets of the Internet, certain giant tech companies' services gained a monopoly or came close to doing so. This process has emerged in connection with gatekeepers of a specific type: the most important online platforms (social media, video sharing, search engines, web stores). In this way, a new, unexpected obstacle to the exercise of freedom of speech appeared, with the result that the earlier constitutional doctrines could no longer be applied without any change. The crux of the problem is that the platforms are privately owned. In formal terms, they are simply market players which are not bound by the guarantees of freedom of speech imposed on public bodies and which may enjoy the protection of freedom of speech themselves.

This Article addresses the issue of the restriction of freedom of speech by social media platforms. Section II delineates traditional media and social media platforms, which is a prerequisite for further reflection on appropriate regulation. Section III examines the issues raised by the deletion of user content by platforms. Based on the fundamentals of European media regulation, Section IV raises the issue of the responsibility of social media platforms to maintain the appropriate quality of democratic publicity. Closing the article, Section V summarizes the conclusions. This Article will present European and U.S. regulatory approaches in parallel, considering both legacy media and social media platforms. The Author of this Article, coming from Europe, undertakes to place the European approach at the forefront, highlighting where it conflicts with the U.S. concept of freedom of speech. However, given that the issues raised by social media platforms are similar everywhere and their regulation is the subject of similar debates worldwide, it cannot be ruled out that European solutions could at least help shape the U.S. academic community's further thinking.

II. DIFFERENTIATION BETWEEN SOCIAL MEDIA PLATFORMS AND LEGACY MEDIA

The operation of social media platforms is fundamentally different from that of "legacy" media. In essence, content on these platforms is created independently of the platforms. However, in the process of publishing and the aftermath, the platform becomes similar in operation to the media and the editorial activity they perform. This fundamental discrepancy and the similarity which exists call for a precise delineation of services in order to define the precise set of liability rules applicable to them. In practice, this means examining

whether existing and mature regulatory solutions and liability regimes for legacy media can be applied, at least in some respects, to social media.

A. The Notion of Online Gatekeepers and Platforms

A gatekeeper is an entity tasked with deciding if a person or thing can pass through a "gate" controlled by the gatekeeper.[1] Gatekeepers have existed in all historic periods of public communication and defining their legal status has often caused problems for the law. Generally, newspaper kiosks, postal carriers, or cable and satellite providers were not considered to have a direct impact on the media content they made available to the public. A postal carrier or cable provider could prevent individual readers or viewers from accessing information by refusing to deliver a paper or fix a network error (thereby also hurting its own financial interests), but it was not in a position to decide on the content of newspaper articles or television programs. Such actors had limited potential to interfere with the communication process, even though they were indispensable parts of it, and this made them a tempting target for governments seeking to regulate, or at least keep within certain boundaries, the freedom of speech of others by regulating the intermediaries.

Even though the Internet seems to provide direct and unconditional access for persons wishing to exercise their freedom of speech in public, gatekeepers still remain an indispensable part of the communication process. A gatekeeper is more specifically defined as a person or entity whose activity is necessary for publishing the opinion of another person or entity, and gatekeepers include Internet service providers, blog host providers, social media, search engine providers, entities selling apps, webstores, news portals, news aggregating sites, and the content providers of websites who can decide on the publication of comments to individual posts. Some gatekeepers may be influential or even indispensable, with a considerable impact on public communication, while other gatekeepers may have more limited powers, and may even go unnoticed by the public. It is true that all gatekeepers are capable of influencing the public without being government actors, and that they are usually even more effective at influencing it than governments themselves.[2] As private entities, they are not bound by the constitutional rules pertaining to free speech, so they can establish their own service rules concerning that freedom.

1. EMILY B. LAIDLAW, REGULATING SPEECH IN CYBERSPACE 37 (2015).
2. *Id.* at 39.

According to the classification developed by Emily Laidlaw, "Internet gatekeepers" form the largest group and they control the flow of information. Among these entities, the "Internet information gatekeepers" form a smaller group, and through this control they are capable of affecting individuals' participation in democratic discourse and public debate.[3] In this model, a gatekeeper belongs to the latter group if it is capable of facilitating or hindering democratic discourse.[4] Such activities raise more direct questions regarding the enforcement of the freedom of speech, on the side both of the party influenced by the gatekeeper and of the gatekeeper itself.

As Uta Kohl notes, the most important theoretical questions pertaining to the gatekeepers of the Internet relate to whether they play an active or a passive role in the communication process, the nature of their editorial activities, and the extent of the similarities between their editorial activities and actual editing.[5] The role of online gatekeepers is usually not passive. They are key actors of the democratic public sphere and actively involved in the communication process, including making decisions about what their users can access and what they cannot, or can access only with substantial difficulty. The European Union (E.U.) Directive, which regulates, in part, the activities of individual gatekeepers, does not require such gatekeepers to acknowledge their own role as editors. But it does allow them to be held liable for infringements in accordance with their relationship with the content. Gatekeepers may not be held liable if they are not actively involved in the public transmission of unlawful content, or if it is not aware of the infringing nature of the content, but they are required to remove such content after becoming aware of the infringement.[6] However, this does not prevent gatekeepers from sorting through the various pieces of content of their own volition and in a manner permitted by law. Under the current legal approach, gatekeepers are not considered as "media services." This means that while they do demand protection for the freedom of speech in order to enable their selection activities, they are not bound by the various legal guarantees

3. *Id.* at 44.

4. *Id.* at 46.

5. Uta Kohl, *Intermediaries within Online Regulation*, *in* INFORMATION TECHNOLOGY LAW 85–87 (Diane Rowland, Uta Kohl & Andrew Charlesworth eds., 5th ed. 2016).

6. 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', 'E-Commerce Directive'), 2000 O.J. (L 178), arts. 12–15.

concerning the right of individuals to access the media.[7] They are also not subject to obligations that are otherwise applicable to the media as a private institution of constitutional value,[8] as it is conceptualized in the European legal approach.[9]

Online platforms are considered among the most influential gatekeepers in the online sphere. The term online platform "refers to an undertaking operating in two (or multi)-sided markets, which uses the Internet to enable interactions between two or more distinct but interdependent groups of users so as to generate value for at least one of the groups."[10] Search engines, new aggregators, online marketplaces, audiovisual and music platforms, video sharing platforms, and social media are all different types of online platforms. Several definitions exist for social media platforms. For example, according to Aleksandra Gebicka and Andreas Heinemann, social media platforms are "web-based services that allow individuals to construct a public or semi-public profile within a limited forum, to articulate a list of other users with whom they share a connection (friends on Facebook), and to view and traverse their list of connections and those made by others within the system."[11] Online platforms are used not only by private individuals but also by commercial enterprises, public figures (politicians among them), and mainstream media outlets, to name a few.

B. Platform Speech and Media Speech

The concept of editorial activities is a key part of the notion of "media."[12] Theoretically, if the activities of gatekeepers are similar to

7. RIGHTS OF ACCESS TO THE MEDIA (András Sajó & Monroe Price eds., 1996).

8. William J. Brennan, *Address,* 32 RUTGERS L. REV. 173 (1979).

9. COMMISSION ON FREEDOM OF THE PRESS, A FREE AND RESPONSIBLE PRESS, A GENERAL REPORT ON MASS COMMUNICATION: NEWSPAPERS, RADIO, MOTION PICTURES, MAGAZINES, AND BOOKS (1947); JOHN C. NERONE, LAST RIGHTS: REVISITING FOUR THEORIES OF THE PRESS 77-100 (1995).

10. European Commission, *Public Consultation on the Regulatory Environment for Platforms, Online Intermediariesm Data and Cloud Computing and the Collaborative*, ECONOMY 5 (2015), https://ec.europa.eu/digital-agenda/en/news/public-consultation-regulatory-environment-platforms-onlineintermediaries-data-and-cloud.

11. Aleksandra Gebicka & Andreas Heinemann, *Social Media & Competition Law*, WORLD COMPETITION 149, 152 (2014).

12. Matthew Ingram, *Sorry Mark Zuckerberg, but Facebook is Definitely a Media Company*, FORTUNE (Aug. 30, 2016), http://fortune.com/2016/08/30/facebook-media-company; Samuel Gibbs, *Mark Zuckerberg Appears to Finally Admit Facebook is a Media Company*, THE GUARDIAN (Dec. 22, 2016), https://www.theguardian.com/technology/2016/dec/22/mark-zuckerberg-appears-to-finally-admit-facebook-is-a-media-company.

such editorial activities, the gatekeepers themselves may be subject to media regulation to a certain extent; otherwise, they may be considered technology companies.

The terms "editing," "editorial decision-making," and "editorial discretion" are usually not defined in legal documents. For the media, these refer to making unavoidable decisions on the content of a given medium, decisions which are indispensable for the operation of any medium. Note that in the context of the press, radio, television, and a considerable number of websites, editors make decisions on content that was commissioned by them or produced by their colleagues like journalists or producers.

The definition of commonplace editorial activity is set in the E.U. Audiovisual Media Services (AVMS) Directive within the following notion of "editorial responsibility":

> 'editorial responsibility' means the exercise of effective control both over the selection of the programs and over their organization either in a chronological schedule, in the case of television broadcasts, or in a catalogue, in the case of on-demand audiovisual media services. Editorial responsibility does not necessarily imply any legal liability under national law for the content or the services provided.[13]

According to this definition, the editor of a media service is the person who selects and compiles the programs to be published, and without whom such programs would not reach the public. The editor of a press outlet commissions and selects the articles to be published in that paper. This traditional editorial control has several components: The editor influences (1) the creation of the content by instructing the journalist, commissions content from external suppliers, (2) the publication of the content, and also (3) how, where, and when the content becomes accessible compared to other content.[14] Editorial control and editorial responsibility are different notions, but in legacy media in general one entails the other: the editor deciding on the publication is responsible for any infringement caused by the content. The activities of social media platforms are considerably different from this model, as the users produce and share independent content en masse, and the platform operator normally does not interfere with this

13. Directive 2010/13/EU on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (AVMS Directive), 2010 O.J. (L 95), art. 1(d)(bb).

14. Max Z. van Drunen, *The Post-Editorial Control Era: How EU media Law Matches Platforms' Organisational Control with Cooperative Responsibility*, 12(2) JOURNAL OF MEDIA LAW 166, 169–71 (2020).

process. However, some aspects of this model are quite similar to traditional editing. Gatekeepers do not usually decide on the publication of any content prior to its publication, but they may decide to remove a piece of content subsequently either voluntarily or to perform a legal obligation. In certain cases, gatekeepers may even prevent the publication of a piece of content through preliminary filtering, and similarly they may display some content to the users in a prominent place while almost hiding other content.

The activities of social media platforms are also characterized by two of the three above-listed elements of editorial control. The platform can make a decision on the publication of the content, although—unlike legacy media—this does not represent a full preliminary decision but, for example, the setting of filters or deletion subsequently. Furthermore, the platform also decides how, where, and to whom user content will be made available. Max van Drunen calls this "organi[z]ational control."[15] It is largely up to the platform to decide (or dependent on the settings of the algorithm regulating this issue) which user may access what content, what appears in a prominent location for him, what content he needs to search for, and what he cannot access at all. In addition, user decisions (namely, which user is a "friend," what they "like" or mark as important, etc.) themselves influence what content the platform offers them. In general, social media platforms can influence the content (for example, the newsfeed of Facebook) displayed to their users in line with their own interests. Notably, such editing is performed in bulk and on a daily basis, using both artificial intelligence and human resources, with a view to improving service quality or serving business or other interests. In Europe, such editing is performed to comply with legal obligations if required for the removal of violating content[16] or the protection of personal data.[17]

It is clear the law is heading towards regarding gatekeepers as editors, and another step in this direction is the 2018 amendment of the AVMS Directive on the regulation of video sharing platforms including social media platforms allowing the publication of audiovisual content. While the AVMS Directive emphasizes that "video-sharing platform services" do not bear any "editorial responsibility," they still may be subject to content-related obligations with regard to the protection of children and taking action against hate speech. The Directive

15. *Id*. at 171–173.

16. E-Commerce Directive, *supra* note 6, arts. 12–14.

17. Case C-131/12, 13 May 2014, (CVRIA, No. Google Spain SL, Google, Inc. v. Agencia Española de Protección Datos Mario Costeja González, judgment [GC], 13 May 2014).

recognizes that such platforms organize (such as display, tag, and sequence) user content, and, when accompanied by an obligation to take action against infringing content, their role is clearly similar to editing:

> (aa) "video-sharing platform service" means a service, as defined by Articles 56 and 57 of the Treaty on the Functioning of the European Union, [which meets the following requirements:] . . .
>
> [(i) the service consists of the storage of a large amount of] program[]es [or] user-generated videos, . . . for which the video-sharing platform provider does not have editorial responsibility;
>
> [(ii) the organi[z]ation of the stored content is determined by the provider of the service including] by automatic means or algorithms, in particular by hosting, displaying, tagging and sequencing.[18]

As Philip Napoli notes, there is nothing new about the media wanting to provide their audience with what they are looking for, what they are interested in, or what they enjoy watching, listening to, or reading. Social media platforms also do this, *par excellence*, mapping user needs with much more sophisticated tools.[19] Platforms decide which user content is accessible (by deleting content which public authorities order to be removed or that violate their own policies) and which of the theoretically available content does actually appear to users. Social media platforms are both indispensable helpers of the legacy media, making masses of people accessible to them, and of their direct competitors competing for commercial revenue coming from the same sources. Large platforms are part of the public sphere as arenas for disseminating news, opinion articles, and information of public interest. Their market position and their ability to intervene in what is offered also make the platforms a *de facto* news service, even if not in its legal sense.[20]

From the regulatory point of view, treating platforms in the same way as legacy media in terms of editorial control would have quite different consequences in Europe and the U.S. The constitutional protection of the right to "press freedom" has led to different conclusions

---

18. Directive 2018/1808/EU of the European Parliament and of the Council of 14 November 2018 amending Directive 2010/13/EU on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive) in view of changing market realities (new AVMS directive), 2018 O.J. (L 303), art. 1(b)(aa).

19. PHILIP M. NAPOLI, SOCIAL MEDIA AND THE PUBLIC INTEREST 11–14 (2019).

20. *Id*. at 13.

on the two sides of the Atlantic.[21] In the United States, the restrictions on freedom of speech and freedom of the press coincide, since in the interpretation of the Constitution and in the case law of the Supreme Court[22]—contrary to Europe—the exercise of freedom of the press does not entail any additional obligation. In the U.S., content regulation is only permitted to a very narrow extent,[23] and editorial freedom cannot be interfered with by state regulations at all.[24] As such, if the U.S. legal system were to regard social media as a news service and the companies engaged in editorial activities as media actors, it would not directly reduce its freedom to assess and arrange user content, but it may also have other consequences, detrimental to the platforms.[25] Conversely, if all this took place in Europe, theoretically, a platform could benefit from and exercise additional rights related to media freedom, but it would also be subject to certain additional obligations.[26] Platforms are not eager to be considered "media," even in the U.S., because it would set a dangerous example that could weaken their European position or have detrimental consequences for them in other countries.

C. The Problem of Artificial Intelligence

Due to the large volumes of data transmitted, gatekeepers use not only human resources, but also algorithms to process information. A term borrowed from mathematics, an algorithm is a method, guideline, or set of instructions that consists of a sequence of steps and is suitable for solving a problem.[27] In general, computer programs embody algorithms used to instruct a computer how to execute a task. In the context of gatekeepers, a decision concerning the flow of information (that is the filtering, removal of higher or lower ranking of content and its presentation to users) is usually determined by an algorithm, meaning that the legal status of such decisions, as well as the nature and subject of legal rights and obligations, poses fundamental

21. LEONARD W. LEVY, EMERGENCE OF A FREE PRESS (2004); STEPHEN M. FEILDMAN, FREE EXPRESSION AND DEMOCRACY IN AMERICA: A HISTORY (2015); David Anderson, *Origins of the Free Press Clause*, 30(3) UCLA L. REV. 455 (1983).

22. Sonja West, *Awakening the Press Clause*, 58 UCLA L. REV. 1025, 1068–1070 (2010–2011); Sonja West, *Press Exceptionalism*, 127(8) HARV. L. REV. 2443 (2014).

23. Federal Communications Commission v. Pacifica Foundation, 438 U.S. 726 (1978).

24. Miami Herald Pub. Co., Div. of Knight Newspapers, Inc. v. Tornillo, 418 U.S. 241 (1974).

25. *See infra* Section IV.

26. *See infra* Section IV.

27. *Definition of Algorithm*, MERRIAM-WEBSTER ONLINE DICTIONARY, https://www.merriam-webster.com/dictionary/algorithm.

questions. An algorithm is a kind of editor, which presents the user with content according to the decisions of its creator and employing data collected about the user during the use of the service or other services (concerning his or her interests and preferences).[28] The control exercised by social media platforms includes the power to decide who may reach the public, who is banned from the public, who is to follow the rules of the public, and who is to remain silent.[29]

Hence, if editing by social media platforms is largely, but not exclusively, carried out by these algorithms, the question rightly arises as to whether this activity can actually be considered to belong within the scope of editorial freedom and whether it needs to consequently also be subject in Europe to restrictions on editorial freedom. Individual legal systems treat the legacy media's editing activities in the same way as speech, and unnecessary or excessive restrictions on editorial freedom are seen as an infringement of freedom of the press. Consequently, the question of whether the outcome of the operation of artificial intelligence and algorithms constitutes editing, and therefore whether this qualifies as speech needs to be answered.

It is a reasonable question whether the communication produced with the help of algorithms used by gatekeepers is protected by the freedom of speech. If an algorithm conducts editing, meaning that it makes decisions concerning the sorting, removal, and ranking of pieces of content, it might be considered speech. Such decisions have a fundamental impact on the public appearance of the actual speaker, who is preferred or disfavored by the algorithm, giving the decision of the algorithm a certain communicative content that is protected under the aegis of the freedom of speech. Such decisions also convey a material communicative message to other users, which influences the capability of such users to access information. For this reason, such users experience the decision as an opinion even more directly.[30] On the other hand, it may be argued that a decision made by an algorithm—such as a search ranking or the compilation of a news stream—is fully automated and without any actual content (it only sorts through or makes other kinds of decisions concerning the content of others), meaning that it should be considered as an action instead of speech. Indeed, the algorithms of gatekeepers often operate without any

28. Sue Halpern, *Mind Control & the Internet*, NEW YORK REVIEW OF BOOKS (June 23, 2011) https://www.nybooks.com/articles/2011/06/23/mind-control-and-internet/?printpage=true.

29. Andrew Tutt, *The New Speech*, 41 CLQ 235, 250 (2014).

30. Stuart M. Benjamin, *Algorithms and Speech*, 161 U. PA. L. REV. 1445, 1447 (2013).

communicative content—think of the operation of the TCP/IP protocol or cache storage, which do not convey any message to users.[31]

Tim Wu argues that the activities of devices needed to convey speech, but which merely transmit information without making any decision, may not be considered speech.[32] A typical embodiment of this proposition in the offline world is a telephone service. On the other hand, cable television services are different, in that cable service providers make decisions or edit how to present a channel to the audience.[33] Wu also argues that the activities of an online gatekeeper should be considered action instead of speech if they are merely functional, in the sense that they are necessary to transmit the speech of others but do not carry any independent meaning themselves and considers the search rankings of a search engine to be without such meaning.[34] The legal approach toward search engines is a complex matter and will be revisited later. Probably even Wu would agree the activities of Facebook go beyond being merely functional and convey material messages in and of themselves, because a personalized news stream is compiled for each user upon login (selecting some of the content available to the user). It seems inevitable that the activities of an editing algorithm must be considered speech, as the algorithm conveys a material message itself. In addition to the trends in legal development, a reason for this is that such activities are experienced by their recipients as speech. However, it seems unlikely that decisions made by algorithms and human beings can be distinguished from each other in a consistent manner.[35]

If they are not recognized as speech, the most important services and activities of social media platforms can be regulated without respect for the most fundamental constitutional protection, that is, the guarantees of freedom of expression. Whereas recognition as speech also implies that the algorithms would also be expected to respect the limits of freedom of speech, and the providers of services using such algorithms would not be exempted from the application of general laws that are not related to the content of speech (such as anti-trust or tax laws).

It also seems clear that such communicating algorithms do not make decisions on their own, and the actual person who created the program

---

31. *Id*. at 1471.

32. Tim Wu, *Machine Speech*, 161 U. PA. L. REV. 1495, 1525–33 (2013).

33. Turner Broadcasting System v. Federal Communications Commission, 512 U.S. 622 (1994); Turner Broadcasting System v. Federal Communications Commission, 520 U.S. 180 (1997).

34. Wu, *supra* note 32, at 1524–31.

35. Benjamin, *supra* note 30, at 1493.

is always there in the background, generally making editorial decisions, that is, determining the way the program should operate. In the end, the "machine does not speak"[36] and, obviously, it does not become a beneficiary of freedom of speech. Still, we can agree with Stuart Benjamin, who stated that, in terms of the exercise of free speech, differentiation between decisions made by the algorithm and by humans can only be done arbitrarily.[37] A decision made by an algorithm (deleting or, on the contrary, highlighting a piece of content, or possibly hiding it without deleting it) carries meaning, therefore constitutes speech for the legal system, so the company operating the algorithm can benefit from the constitutional protection of speech, but must also bear its burden. Accepting this would involve legal systems, for example, in the case of Facebook, identifying each user feed compiled by the platform's algorithm, the compilation of which is influenced by the user's decisions as speech, and defining the outcome of editorial activity, and the platform running the algorithm as a publisher in legal terms, together with all the inherent positive and negative consequences that come with it.[38] After all, the platform itself has the final say on what is published on the platform; it retains editorial control, even if such control differs from editing in legacy media in many respects.

III. THE REGULATION OF FREEDOM OF EXPRESSION BY SOCIAL MEDIA PLATFORMS

Platforms are not neutral actors in the public communication process. Their decisions on content selection are protected expressions of the platforms' freedom of speech. At the same time, similarly to legacy media, legislation also limits the decision-making freedom of platforms, imposing obligations on them to take action against illegal content. The nature of this obligation and the liability regime associated with it differ fundamentally on the two sides of the Atlantic. In addition, user content may be deleted, filtered, or blocked at the discretion of the platform, in accordance with the terms of the contract it concluded with the users. The rules of the two parallel sets of norms, that is those of the legal regulation by each system, are also binding on platforms and those of the contract, and apply only between the platform and the user. These differ by protecting or restricting users'

36. *Id.* at 1479.

37. *Id.* at 1493–94; Stuart M. Benjamin, *The First Amendment and Algorithms*, THE CAMBRIDGE HANDBOOK OF THE LAW OF ALGORITHMS 630–31 (Woodrow Barfield ed., 2021).

38. Alan M. Sears, *Algorithmic Speech and Freedom of Expression*, 53(4) VANDERBILT JOURNAL OF TRANSNATIONAL LAW 1327, 1373 (2020).

freedom of speech to different degrees. This discrepancy brings detrimental and dangerous consequences for the public sphere.

A. The Regulation of Social Media Platforms in Europe and in the United States—A Quick Overview

1. The European Union and its Member States

If gatekeepers provide only technical services when they make available, store, or transmit the content of others (much like a printing house or a newspaper stand), then it would seem unjustified to hold them liable for the violations of others, as long as they are unaware of the violation. However, according to the European approach, the gatekeepers may be held liable for their own failure to act after becoming aware of the violation if they fail to remove the infringing material. The relevant E.U. Directive requires intermediaries to remove such materials after becoming aware of their infringing nature.[39]

Even though the gatekeeper activities falling within the scope of the Directive (namely, mere conduit, caching, and hosting) play an important role in online communication, the issue of liability has arisen since 2000, the year of adoption of the Directive, with regard to various gatekeepers that did not even exist at the time, or which were not included in the scope of legislation for other reasons (for example, search engines or social media platforms). In the absence of a better analogy, courts tend to liken such entities, for example, to hosting providers. The material scope of the regulation is of great importance: if certain conditions are met, the E.U. Directive exempts gatekeepers from liability even if they let violating pieces of content pass. This exemption-based system should not necessarily be considered outdated, but something has certainly changed since 2000: there are fewer and fewer reasons to believe that the gatekeepers of today remain passive regarding content and perform nothing more than storage and transmission. While content is still produced by users or other independent actors, the services of gatekeepers select from and organize, promote, or reduce the ranking of such content, and may even delete it or make it unavailable within the system. The fairness rule of the Directive that grants exemption to a passive actor as long as it does not get involved in the process (that is, until it becomes aware of the violation) seems less and less to be the only feasible solution to the problem of the new gatekeepers. Still, it seems true that the volume of content processed by the new gatekeepers makes it impossible and unreasonable to impose a comprehensive obligation to control prior to

39. E-Commerce Directive, *supra* note 6, arts. 12–14.

publication, or even after publication without the requirement of an external notice. Accordingly, Articles 12 to 14 grant wide exemptions to gatekeepers. For hosting providers (social media platforms among them), this means that a provider is not held liable for the transmission or storage of infringing content, if it is not its own content and it is not aware of the infringing nature of the content, provided that it takes action to remove the information or make it inaccessible without delay.[40] If such measures are not taken, the provider may be held liable for its own omission. In addition, the Directive also stipulates that intermediaries may not be subject to a general monitoring obligation to identify illegal activities.[41]

The notion of "illegal" raises an important issue, as the obligation of removal is independent from the outcome of an eventual court or official procedure that may establish the violation, and the storage provider is required to take action before a decision is passed, provided that a legal procedure is initiated at all. This means that the provider has to decide on the issue of illegality on its own, and its decision is free from any legal guarantee, even though it may have an impact on freedom of speech. This rule may encourage the provider concerned to remove content to escape liability, even in highly questionable situations. It would be comforting—but probably inadequate, considering the speed of communication—if the liability of an intermediary could not be established unless the illegal nature of the content it has not removed is established by a court.[42]

The interpretation of the Directive has been somewhat clarified by certain decisions of the Court of Justice of the European Union. One of the cases related to the protection of intellectual property. In *Google France v. Louis Vuitton*,[43] the Court held that the storage provider (Google's AdWords service) was exempted from liability, as:

> "the rule laid down therein applies to an Internet referencing service provider in the case where that service provider has not played an active role of such a kind as to give it knowledge of, or control over, the data stored. If it has not played such a role, that service provider cannot be held liable for the data which it has stored at the request of

40. *Id*. at art. 14.

41. *Id*. at art. 15.

42. Christina M. Mulligan, *Technological Intermediaries and Freedom of the Press*, 66 SMU L. REV. 157, 175 (2013).

43. Joined Cases of C-236/08 to C-238/08, Google France S.A.R.L. and Google, Inc. v. Louis Vuitton Malletier S.A. and Others, ECLI:EU:C:2010:159. Cour de Cassation [Final court of appeals] Joined Cases of C-236/08 to C-238/08, Google France S.A.R.L. and Google, Inc. v. Louis Vuitton Malletier S.A. and Others, 23 March 2010.

> an advertiser," provided that it took action after becoming aware of the situation.[44]

According to the judgment handed down in *L'Oréal SA and Others v. eBay International and Others*,[45] the operator of a webstore is not liable for any content uploaded by a client because "the mere fact that the operator of an online marketplace stores offers for sale on its server, sets the terms of its service, is remunerated for that service and provides general information to its customers cannot have the effect of denying it the exemptions from liability."[46] However, the exemption applies only as long as it remains neutral toward the respective piece of content:

> Where, by contrast, the operator has provided assistance which entails, in particular, optimi[z]ing the presentation of the offers for sale in question or promoting those offers, it must be considered not to have taken a neutral position between the customer-seller concerned and potential buyers but to have played an active role of such a kind as to give it knowledge of, or control over, the data relating to those offers for sale. It cannot then rely, in the case of those data, on the exemption from liability referred to in Article 14(1) of Directive 2000/31.[47]

The question is thus whether the host provider could have been aware of the infringing content.[48] The above rulings are to be regarded in light of the absence of a general monitoring and control obligation, as providers may not be required to implement such general technical solutions (filtering).[49]

In other areas, the situation of gatekeepers seems less comfortable. The directive on combating terrorism sets forth a similar obligation to remove content.[50] Data protection regulations require gatekeepers to

---

44. *Id.* at para. 120.

45. Case C-324/09, L'Oréal S.A. and Others v. eBay International A.G. and Others, ECLI:EU:C:2011:474.

46. *Id.* at para. 115.

47. *Id.* at para. 116.

48. *Id.* at para. 120.

49. *See also* Case C-70/10, Scarlet Extended S.A. v. Société belge des auteurs, compositeurs et éditeurs S.C.R.L. (SABAM), ECLI:EU:C:2011:771; Case C-360/10 Belgische Vereniging van Auteurs, Componisten en Uitgevers C.V.B.A. (SABAM) v. Netlog N.V., ECLI:EU:C:2012:85.

50. Directive (EU) 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, 2017 O.J. (L 88), Art. 21.

comply with data protection rules regarding their own activities (that is, the obligation does not arise as a result of a possible violation by another person or user).[51] It also seems possible for a court of a member state to decide that the activities of a certain gatekeeper are not covered by the E-Commerce Directive (namely, it does not qualify as a host provider), meaning that the general rules of civil and criminal law may be applied. The European Court of Human Rights (ECtHR) did not object to this interpretation when the content provider of a website was held liable by a national court for comments posted anonymously.[52]

On-demand media services have fallen within the scope of the AVMS Directive since 2007. Such services can be accessed through the Internet, but social media are not one form of such services. The main reason for this is that providers of on-demand media services bear editorial responsibility for the content they publish, they order and purchase such content, and they have a final say in publishing a piece of content.[53] However, social media only provides a communication platform because it may not make any decision regarding a piece of content before it is published. The situation is different if some kind of preliminary filtering is used, but such filtering affects only specific types of content.

As social media platforms spread, it became clear, about a decade after the previous amendment of the Directive, that media regulation could not be interpreted in such a restrictive manner any longer. As already mentioned, the recently amended AVMS Directive introduced the terms "video-sharing platform service" and "video-sharing platform provider."[54] According to the amendment eventually adopted in November 2018, the material scope of the Directive was extended to cover such services. Even though the original proposal would not have extended the scope of the Directive to social media platforms (in terms of the audiovisual content uploaded to the site), it became clear during the legislative process that they could not be exempted from the Directive, and it could not focus solely on portals used to share videos,

---

See Monica Horten, *Content "Responsibility": The Looming Cloud of Uncertainty for Internet Intermediaries* 12, CENTER FOR DEMOCRACY AND TECHNOLOGY.

51. Case C-131/12 *Google Spain*; Regulation (EU) 2016/679, General Data Protection Regulation (GDPR), art. 17.

52. Eur. Ct. H.R. Delfi A.S. v. Estonia, no. 64569/09, judgment of 15 June 2015 [GC]; Magyar Tartalomszolgáltatók Egyesülete and Index.hu Zrt. v. Hungary, no. 22947/13, judgment of 6 February 2016.

53. AVMS Directive, *supra* note 13, art. 1.

54. New AVMS Directive, *supra* note 18, art. 1(1)b(aa).

such as YouTube.[55] For this reason, the recital of the amending Directive provides:

> Video-sharing platform services provide audiovisual content which is increasingly accessed by the general public, in particular by young people. This is also true with regard to social media services, which have become an important medium to share information and to entertain and educate, including by providing access to programmes and user-generated videos. Those social media services need to be included in the scope of Directive 2010/13/EU because they compete for the same audiences and revenues as audiovisual media services. Furthermore, they also have a considerable impact in that they facilitate the possibility for users to shape and influence the opinions of other users. Therefore, in order to protect minors from harmful content and all citizens from incitement to hatred, violence and terrorism, those services should be covered by Directive 2010/13/EU to the extent that they meet the definition of a video-sharing platform service.[56]

This means that, despite their somewhat misleading name, video-sharing platforms include audiovisual content published on social media platforms. Article 28b of the amended AVMS Directive provides that Articles 12 to 15 of the E-Commerce Directive—in particular, the provisions on hosting service providers and the prohibition on introducing a general monitoring obligation—remain applicable. Member States must ensure that video-sharing platform providers operating within their respective jurisdiction "take appropriate measures" to ensure four things. First, the protection of minors from programs, user-generated videos, and audiovisual commercial communications that may impair their physical, mental, or moral development. Second, the protection of the general public from programs including user-generated videos and audiovisual commercial communications containing incitement to violence or hatred directed against a group of persons or a member of a group. Third, the protection of the general public from programs, user-generated videos, and audiovisual commercial communications containing content the dissemination of which constitutes an activity that is a criminal offense under E.U. law, namely public provocation to commit a terrorist offence within the meaning of Article 5 of Directive (EU) 2017/541, offenses concerning child pornography within the meaning of Article 5(4) of

55. Duncan Robinson, *Social Networks Face Tougher EU Oversight on Video Content* 8, FINANCIAL TIMES (May 23, 2017), https://www.ft.com/content/d5746e06-3fd7-11e7-82b6-896b95f30f58.

56. New AVMS Directive, *supra* note 18, recital (4).

Directive 2011/93/EU of the European Parliament and of the Council, and offen[s]es concerning racism and xenophobia within the meaning of Article 1 of Framework Decision 2008/913/JHA. Fourth, compliance with the requirements set out in Article 9(1) with respect to audiovisual commercial communications that are marketed, sold, or arranged by the video-sharing platform providers (general restrictions of commercial communications and provisions in order to safeguard minors from commercials).

What constitutes an "appropriate measure" is to be determined with regard to the nature of the content in question, the harm it may cause and the characteristics of the category of persons to be protected, as well as the rights and legitimate interests at stake. This includes those of the video-sharing platform providers and the users who created, transmitted, and/or uploaded the content, as well as the public interest.[57] According to the Directive, such measures should extend to eight actions (among others). First, defining and applying the above-mentioned requirements in the terms and conditions of the video-sharing platform providers. Second, establishing and operating transparent and user-friendly mechanisms for users of video-sharing platforms to report or flag up content to the video-sharing platform provider concerned. Third, establishing and operating age verification systems for users of video-sharing platforms with respect to content that may impair the physical, mental, or moral development of minors with a view to protecting children. Fourth, providing parental control systems with respect to content that may be harmful for minors. Fifth, providing users with easy-to-use means of identifying violating content. Sixth, establishing and operating transparent, easy-to-use and efficient procedures for managing and settling disputes between video-sharing platform providers and users. Seventh, providing information and explanations from service providers regarding the protective measures. Eighth, implementing effective measures and controls aimed at media awareness, and providing users with information regarding such measures and controls.[58]

While the new provisions of the Directive appear rather detailed, the major platform providers have already been making efforts to comply with the requirements that have now become mandatory. The regulation applies to only a narrow range of content—namely, audiovisual content—and government is only granted control over the operation of platform providers in connection with a handful of content-related issues, such as protection of minors, hate speech,

57. *Id.* at art. 28b(3).

58. *Id.*

support for terrorism, child pornography, and denial of genocide. Such content is in any case commonly banned or removed by the platforms upon receiving notice of it under their own policies. Nonetheless, not all content prohibited in Europe is inconsistent with such policies. Once the provisions of the Directive are transposed into the national law of E.U. Member States, platform providers will be required to take action under both the E-Commerce Directive and the AVMS Directive. These two pieces of legislation act mostly in parallel, as the former requires infringing content to be removed in general, while the latter defines certain types of infringing content and lays down detailed rules for their removal. The AVMS Directive lays down numerous provisions that both facilitate the application of the rules and work as procedural safeguards.

2. The United States

In the U.S., the liability of gatekeepers is regulated on the basis of a different theoretical background. For the purpose of ensuring the smooth growth and economic strengthening of Internet companies, the courts took a step backwards by claiming to protect freedom of speech.[59] Today, gatekeepers are granted virtually complete immunity when it comes to infringing content of others.

Section 230 of the 1996 Communications Decency Act (CDA)[60] lays down the "Good Samaritan" protection for the providers of "interactive computer services":

> (c) Protection for "Good Samaritan" blocking and screening of offensive material
>
> (1) Treatment of publisher or speaker
>
> 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.
>
> (2) Civil liability
>
> No provider or user of an interactive computer service shall be held liable on account of—

59. Anupam Chander, *How Law Made Silicon Valley,* 63 EMORY LAW JOURNAL 641 (2014).

60. 47 U.S.C. § 230 (2021).

> (A) any action voluntarily taken in good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected; or
>
> (B) any action taken to enable or make available to information content providers or others the technical means to restrict access to material described in paragraph (1).[61]

However, the protection is not complete and unconditional. The provision relies on judicial case law to establish when a gatekeeper becomes a publisher or speaker, thereby losing its immunity, and Subparagraph (E) stipulates that the protection does not apply in the event of committing a federal crime, or violating communications privacy laws, sex trafficking laws, or intellectual property rights.

It seems clear that the CDA explicitly allows gatekeepers to make content-related decisions: gatekeepers are free to decide which pieces of content to remove and how to present content to their users, and this fact implies the restriction of their freedom of speech, even if not in the legal sense.[62] Section 230 has given rise to extensive case law, which seems to interpret the obligations of a "Good Samaritan" in a restrictive manner. In *Zeran v. AOL*,[63] the court established that exemption from liability does not cease to exist when a letter or takedown notice is sent to a service provider drawing its attention to the infringing content. The judgment also noted that the publication, editing and removal of a piece of content falls within the discretion of the service provider, and it does not exclude its immunity.

In *Fair Housing Council of San Fernando Valley v. Roommates.com*,[64] the United States Court of Appeals for the Ninth Circuit ruled against the operator of the website. The website was meant to bring together co-tenants (such as students living away from home), so that accommodation-seekers could specify the features of persons with whom they would be willing to live, thereby excluding persons of different races and colors, while they were required to specify their own characteristics. The website operator was held liable for such discriminatory practices, as it encouraged its users to violate the requirement of equal treatment by performing targeted searches.

---

61. *Id.*

62. *See* Rebecca Tushnet, *Power Without Responsibility: Intermediaries and the First Amendment*, 76 GEO. WASH. L. REV. 986, 1009 (2008).

63. 129 F.3d 327 (1997).

64. 521 F.3d 1157 (2008).

This interpretation of the law was weakened by the judgment passed in *Jane Doe v. Backpage.com*.[65] The users of the website in this case were allowed to publish advertisements, including solicitations for prostitution, and the plaintiffs argued that this may have facilitated the trafficking of human beings for sexual purposes. The United States Court of Appeals for the First Circuit held that the decision made by the website operator on the organization of user content and establishing the rules of publication did not prevent him from invoking the immunity granted by section 230 CDA. This decision goes against the conclusions reached in the above-mentioned *Roommates.com* case. Backpage's offering of adult services sections remained highly controversial, due to allegations that Backpage knowingly allowed and encouraged users to post ads relating to prostitution and human trafficking, particularly involving minors, and took steps to intentionally obfuscate the activities. After a series of court cases and the arrest of the company's CEO and other officials, Backpage removed the adult services subsection in the U.S. in 2017. On April 6, 2018, Backpage.com and affiliated websites were seized by the federal law enforcement bodies.[66]

Doe v. MySpace, Inc.,[67] involved a social media service through which a user contacted and, in the course of a personal meeting, eventually sexually molested another user. The United States Court of Appeals for the Fifth Circuitt held that the website provided only a means of communication to its users, and it was not held liable for the crime committed. Similarly, a service provider may not be held liable if its social media service is used to organize the perpetration of an act of terror.[68]

B. Protecting User Speech from Social Media Platforms

1. Private Regulation by Social Media Platforms—A Primer

Under the European regulation, social media platforms can be forced to assume some kinds of editorial tasks, as the law requires them to assess the legality of content and to remove illegal content when notified. There are other situations where platforms proceed on their

65. 817 F.3d 12 (2016).

66. Paul Demko, *The Sex-Trafficking Case Testing the Limits of the First Amendment*, POLITICO (July 29, 2018), https://www.politico.com/magazine/story/2018/07/29/first-amendment-limits-backpage-escort-ads-219034.

67. 528 F.3d 413 (5th Cir. 2008).

68. Fields v. Twitter, Inc., 217 F. Supp. 3d 1116 (N.D. Cal. 2016); Cohen v. Facebook, Inc., 252 F. Supp. 3d 140 (E.D.N.Y. 2017).

own initiative and decided on the status of user-generated content. Jack Balkin calls this phenomenon "private governance."[69] Others prefer to use the less euphemistic term "private censorship."[70] However, the term "private regulation" also seems to capture the essence of the matter, whereby a platform provider influences the publication or further accessibility of content published by users to an extent and in a manner permitted by law by exercising its ownership rights over the platform and other rights stipulated in its contract with the users.

Platform providers can have different motives for adopting private regulations. An obvious motive for doing so is to protect their business interests. Platform providers have an interest in making sure that their users feel safe while using their platform and are not confronted by insulting, upsetting, or disturbing content. The moderation and removal of such content is not done in line with the limitations of free speech, meaning that a piece of content may be removed using this logic even if it would otherwise be permitted by law, while a piece of content may remain available even if it violates the limitations of free speech. The typically U.S.–owned and established platforms are in a strange and somewhat ambivalent situation. On the one hand, their activities are protected by the First Amendment and the CDA, and their developers and employees represent a culture of American–style free speech. On the other hand, the private regulation they apply provides far less protection for public speech than the U.S. legal system.[71] Moreover, Facebook also tends to remove pieces of content that are clearly protected by the freedom of speech in Europe in an attempt to provide a "safe space" for its users.[72]

A major problem with private regulation is that it may be more strict and more lenient than government regulation, and, as a result, the regulation of content is unpredictable. Another significant concern is that there is no adequate decision-making procedure in place regarding

---

69. Jack Balkin, *Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation*, 51 U.C. DAVIS L. REV. 1179, 1182 (2018).

70. Marjorie Heins, *The Brave New World of Social Media Censorship*, 127 HARV. L. REV. FORUM 325, 325 (2014).

71. Balkin, *supra* note 69, at 1195; Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1599, 1625 (2018).

72. *See*, *e.g.*, Cecilia Rodriguez, *Facebook Finally Lands in French Court for Deleting Nude Courbet Painting*, FORBES (Feb. 5, 2019), https://www.forbes.com/sites/ceciliarodriguez/2018/02/05/facebook-finally-lands-in-french-court-for-deleting-nude-courbet-painting (some questionable editorial decisions made based on the general prohibition of nudity); Sam Levin, Julia C Wong, & Luke Harding, *Facebook Backs Down from "Napalm Girl" Censorship and Reinstates Photo*, THE GUARDIAN (Sept. 9, 2016), https://www.theguardian.com/technology/2016/sep/09/facebook-reinstates-napalm-girl-photo.

the removal of pieces of content, meaning that the constitutional safeguards commonly available in legal proceedings (such as notification of users concerned, possibility of appeal, due process, the identification of the decision-maker, publishing decisions in writing so that they might be known, and so forth) are absent. The absence of an appropriate procedure greatly contributes to the lack of transparency regarding decisions made based on private regulation and does nothing to clarify existing uncertainties concerning the rules applied in such important forums of public life.

In addition to the ownership of a platform, a contract by and between the platform and each user serves as the legal basis for the platform's capacity to interfere with the freedom of speech of its users. The provisions of that contract are determined solely by the platform. Users are not in a position to request the amendment of the contract, while it may be amended by the platform unilaterally at any time. It is also important that the same contract is concluded with each and every user. Even though the contract, and the interference permitted by it, affects the exercise of a constitutional right, and countless debates, conversations, and exchanges of information on public affairs are taking place on the platform at any given time, no interference by the platform can be considered as state action, and the platform itself is not considered a public forum. An action taken by a platform, even if it limits the opinions of its users, cannot be attributed to the government, meaning that it is not subject to any constitutional safeguard relating to the freedom of speech.[73]

In practical terms, the solution to any conflict or dispute that may arise between a platform provider and a user concerning free speech is to be found among the rules of contract law and not the various principles of constitutional law.[74] When a user subscribes to a platform and accepts the terms and conditions of that platform by a simple click of a mouse, he or she becomes subject to "private regulation," including all content-related provisions as well, and the safeguards of free speech are no longer applicable concerning the user's relationship with the platform.[75] It should not come as a surprise that the contracts used by all major platforms are carefully considered and precisely drafted documents (or, conversely, that they use vague language for the very purpose of extending the discretionary powers of the platform). A comparative analysis prepared by Michael Rustad and Thomas Koenig

73. Jacquelyn E. Fradette, *Online Terms of Service: A Shield for First Amendment Scrutiny of Government Action*, 89 NOTRE DAME L. REV. 947, 953–57 (2014).

74. *Id.* at 971.

75. *Id.* at 977.

provides a detailed overview of such contract terms and conditions.[76] Their investigations pointed out numerous concerns pertaining to consumer protection, including the difficulty of reading the provisions, the arbitration clauses used in such contracts that make it difficult for users to file a lawsuit, the vague meaning of various provisions, and so forth.

The current legal framework does not provide users with any powerful means should they find themselves in a quarrel with the platform. Even though Section 230 CDA incentivizes platforms not to use private regulation by granting them immunity regarding illegal content available on the platforms, it certainly does not prohibit private censorship.[77] Moreover, the European concept of the liability of host providers (as adopted pursuant to Article 14 of the E-Commerce Directive) is a direct incentive for platforms to implement private censorship. Regarding the lack of a balance of power between service providers and users, any dispute that may arise between them regarding the enforcement of their contract may be settled within the legal framework of consumer protection.[78] However, this option is available only if the user concerned qualifies as a consumer, meaning that it is not available to institutional users (namely, media businesses).[79]

Furthermore, consumer protection does not seem to provide any broad possibilities for protecting the freedom of speech of users when the platform's policies and their application are reasonable and justifiable but not arbitrary, which they typically are. Even though they might be questionable, it does not suggest any violation of the consumers' rights in and of themselves. It seems also difficult to object to the application of such policies on a legal basis, considering that a platform is free to determine its own policies and instruct its moderators without being required to respect the constitutional safeguards and legal limitations of the freedom of speech. The only option for a user is to show that the platform removed a piece of content it was not authorized to remove[80]—something that seems nothing short of impossible to demonstrate due to the widely defined limitations of

76. Michael L. Rustad & Thomas H. *Koenig, Wolves of the World Wide Web: Reforming Social Networks' Contracting Practices,* 49 WAKE FOREST L. REV. 1431 (2014).

77. Heins, *supra* note 70, at 328.

78. *See* Kevin Park, *Facebook Used Takedown and it was Super Effective! Finding a Framework for Protecting User Rights of Expression on Social Networking Sites*, 68 N.Y.U. ANN. SUR. AM. L. 891 (2013).

79. Supreme Court, Austria, Case C-498/16, Maximilian Schrems v. Facebook Ireland, Ltd., judgment of 25 January 2018.

80. Fradette, *supra* note 73, at 957.

content and the broad discretionary powers of the platform. A user may also try to make use of the existing anti–discrimination rules if his or her right to equal treatment is violated, but producing adequate evidence in such a situation (showing that a piece of content was removed from one user but was not removed when published by another user) seems rather difficult, and the enormous volume of content and the absence of a monitoring obligation on the side of the platform (which may be invoked as a defense by the platform) also considerably limit the chances of a user.

The moderation of user–generated content by a platform interferes with the free speech of its users. Platforms that decide to moderate such content are trying to walk a tightrope between the "chaos of too much freedom" and the "sterility of too much control."[81] Not surprisingly, balancing is not exactly easy. Platforms might be pressured by governments into removing content that is not necessarily illegal without conducting an adequate procedure for a number of reasons.[82] Platforms also have a number of reasons of their own for interfering with their users' free speech. The primary reason, as already mentioned, is the protection of their own business interests by way of filtering and removing content that might scare away other users or any major business partner or advertiser of the platform.

As Kate Klonick pointed out, a social media platform, in the absence of a more appropriate analogy, must be considered as a new and independent regulator (governor). It establishes, controls, and operates its own infrastructure, which is used by users for communication according to its own interests. It also has a centralized organization that follows its own pre-determined rules (even if those are not necessarily accessible to outsiders in detail) and makes *ex ante* or *ex post* decisions regarding various pieces of content.[83] In other words, a platform decides on pieces of content using a particular aggregational theory of free speech. It seeks to become and remain open and attractive for as many users as possible while trying to protect its users from insults or other forms of communication that could scare them away.[84] This strange, aggregated, and hybrid system brings together the principles of the First Amendment and the European approach to free

81. James Grimmelmann, *The Virtues of Moderation*, 17 YALE J. L. & TECH 42 (2015).

82. Benjamin F. Jackson, *Censorship and Freedom of Expression in the Age of Facebook*, 44 N.M. L. REV. 121, 127–29 (2014).

83. Klonick, *supra* note 71, at 1662–64.

84. Brett J. Johnson, *Facebook's Free Speech Balancing Act: Corporate Social Responsibility and Norms of Online Discourse*, 5 U. BALT. J. MED. L. & ETHICS. 19, 33–34 (2016).

speech, all interpreted and applied according to the interests of the platform itself, with possible differences in each state or region (according to the respective territory's government's approach toward free speech and the platform's free activities), through decision–making procedures that are not transparent to the parties concerned.

2. Legislation and Recent Proposals Aiming to Limit the Powers of Social Media to Restrict User Speech

Some European legislatures already consider the obligation of removal set forth in Article 14 of the E-Commerce Directive to be insufficient, and they impose additional obligations on platform providers. The corresponding Act in German law (effective as of January 1, 2018) is a paramount example of this trend.[85] According to the applicable provisions, all platform providers within the scope of the Act (namely, platform providers with over two million users from Germany) must remove all user content that commits certain criminal offences specified by the Act. Such offenses include defamation, incitement to hatred, denial of the Holocaust and the spreading of scare–news.[86] Manifestly unlawful pieces of content must be removed within twenty-four hours after receipt of a notice, while any "ordinary" unlawful content must be removed within seven days.[87] If a platform fails to remove a given piece of content, it may be subject to a fine of up to fifty million euro (theoretically, in cases of severe and multiple violations).[88]

Some argue that this regulation is inconsistent with the E-Commerce Directive, as it provides for a general exception, instead of ad hoc exceptions, from the free movement of services. In addition, the Directive requires urgency as a condition of applying the exception, but the German Act does not refer to specific pieces of content, meaning that it cannot meet that requirement.[89] This piece of German legislation has been widely criticized as limiting the freedom of speech,[90] even

85. Act to Improve Enforcement of the Law in Social Networks (Network Enforcement Act, 2017), [Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken (Netzwerkdurchsetzungsgesetz) Artikel 1 G. v 01.09.2017 BGBl. I S. 3352 (Nr. 61)].

86. *Id.* at Section 1.

87. *Id.* at Section 3.

88. *Id*. at Section 4.

89. Gerald Spindler, *Internet Intermediary Liability Reloaded: The New German Act on Responsibility of Social Networks and its (In-)Compatibility with European Law*, 8 JIPITEC 166, 167–70 (2017) (last visited Feb. 19, 2022).

90. *See*, *e.g.*, *Germany: Flawed Social Media Law*, HUMAN RIGHTS WATCH (Feb. 14, 2018), https://www.hrw.org/news/2018/02/14/germany-flawed-social-media-law.

though it does not go much further than the E.U. Directive itself; it simply refines the provisions of the Directive, lays down the applicable procedural rules, and sets harsh sanctions for violating platforms. Nonetheless, the rules are followed in practice, and Facebook seems eager to perform its obligation to remove objectionable content.[91] The German regulation shows how difficult it is to apply general pieces of legislation and platform–specific rules simultaneously, and it demonstrates how governments seek to have social media platforms act as judges of user-generated content.

France has adopted regulation similar to the German legislation: platforms must quickly remove pieces of content that are incitements to hatred once they have been notified of them.[92] The law adopted on May 13, 2020 (loi Avia) restricted the general rules of the notification and removal procedure in several aspects. Platforms are obliged to delete content that supports terrorism or displays child pornography within one hour of becoming aware of it, and platform providers are obliged to remove content that qualifies as other criminal activity within twenty-four hours. In addition, the law introduced a number of other rules for the management of user content (platforms must set up an efficient and easy-to-use notification system; confirmation of the notification must be sent to the notifier; if the request in the notification is granted, the deletion must be completed within twenty-four hours; a remedy must be provided; and if a piece of content is removed, the author of the content must also be notified if possible, providing him with a remedy, and so forth).[93]

In a decision of the French Constitutional Council (*Conseil Constitutionnel*) on June 18, 2020, several provisions of the loi Avia adopted by the parliament were found to be unconstitutional and were annulled.[94] The Constitutional Council found that freedom of speech may be restricted, but the restrictions must be necessary and proportionate to the objective pursued. According to the Constitutional

---

91. Reuters, *Facebook Deletes Hundreds of Posts Under German Hate-Speech Law*, REUTERS (July 27, 2018), https://uk.reuters.com/article/us-facebook-germany/facebook-deletes-hundreds-of-posts-under-german-hate-speech-law-idUKKBN1KH21L.

92. *France Online Hate Speech Law to Force Social Media Sites to Act Quickly*, THE GUARDIAN (July 9, 2019), https://www.theguardian.com/world/2019/jul/09/france-online-hate-speech-law-social-media.

93. *France: Analysis of Draft Hate Speech Bill*, ARTICLE 19 (July 3, 2019), https://www.article19.org/resources/france-analysis-of-draft-hate-speech-bill;
Avia Bill EN, 2019 (No. 310) (Fr.), https://ecnl.org/sites/default/files/files/Text-of-Avia-Bill-EN.pdf.

94. Décision n° 2020-801 (June 18, 2020), https://www.conseil-constitutionnel.fr/decision/2020/2020801DC.htm.

Council, the infringing nature of the content in question is not self-evident. Its assessment depends solely on the designated authority. Furthermore, an appeal against a decision has no suspensive effect on enforcement and the one-hour timeframe is insufficient for a judicial decision to be made on the matter. Moreover, given the expected high number of notifications, the unspecified number of applicants, and the lack of a judicial decision prior to notification, it is not viable that the service provider will be able to carry out complex legal analyses under the French Penal Code posing the risk of service providers automatically granting requests, in several cases in violation of the freedom of expression. Finally, the sanctions envisaged are so significant that such a restriction on freedom of expression is unconstitutional.

Austria also introduced obligations for online platforms,[95] following the German pattern. The Austrian rules, which have been in force since January 1, 2021, apply to domestic and foreign platforms that have more than 100,000 users in Austria or which have revenues in Austria of more than 500,000 euro. The video content on video-sharing platforms is governed by the rules of the Audiovisual Media Services Act, while the Communication Platforms Act applies to the rest of their content. The new rules require platforms to set up an effective and transparent procedure for reporting and deleting illegal content. Deletion must take place within twenty-four hours if the illegality is "obvious to a legal layman," or within seven days if a detailed examination is necessary. There must be a complaints procedure in place for users affected by deletion or blocking to avoid "overblocking." Failure to comply with these obligations may result in fines of a maximum of ten million euro being imposed on the platform.[96]

Just a few months after President Trump's ban from social media platforms, the Florida state legislature passed a bill that would have banned the suspension of social media accounts of candidates for public office, subject to heavy fines. In addition, it would have allowed the deletion of user content, its "shadowing" (hiding some user content without deletion), or the suspension and deletion of user accounts only subject to strict obligations.[97]

---

95. Communication Platforms Act, 2021 (Austria), https://www.ris.bka.gv.at/Dokumente/Erv/ERV_2020_1_151/ERV_2020_1_151.pdf.

96. *The new legislation against Online Hate Speech—A brief overview*, MGLP (Apr. 26, 2021), https://www.mglp.eu/en/the-new-legislation-against-online-hate-speech-a-brief-overview.

97. H.B. 7013, 2021 Leg., Reg. Sess. (Fl. 2021), *see* https://www.flsenate.gov/Session/Bill/2021/7013/BillText/Filed/PDF.

Industry groups sued the state to overturn the Act a few days after its governor signed it, claiming it violated those companies' First and Fourteenth Amendment rights and that content moderation was allowed under Section 230 CDA. The United States District Court for the Northern District of Florida granted the plaintiff's request for a preliminary injunction, saying that the law was "an effort to rein in social media providers deemed too large and too liberal" and "not a legitimate government interest." It was also deemed discriminatory and potentially violated the First Amendment free speech rights of Big Tech platforms, as it did not apply to smaller platforms, or any platforms owned by a company with a theme park in Florida. Finally, the court stated that "the legislation does not survive strict scrutiny. Parts also are expressly pre-empted by federal law."[98] It was clearly incompatible with Section 230 CDA, which allows platforms to moderate content.[99]

The U.K. has also drafted a new law tailored to online platforms, requiring a duty of care from platforms.[100] The draft Online Safety Bill was published on May 12, 2021.[101] If enacted, the Bill would impose duties of care on providers of online content-sharing platforms and search engines. Ofcom, the U.K. communications authority, would enforce compliance and its powers would include being able to fine companies up to eighteen million pounds or ten percent of their annual global turnover, whichever is higher, and the power to block access to sites. Companies under the scope of the Bill would need to take "robust action to tackle illegal abuse, including swift and effective action against hate crimes, harassment and threats directed at individuals and keep their promises to users about their standards."[102]

"The largest and most popular social media sites (Category 1 services) [would also] need to act on content that is lawful but still

98. Netchoice, LLC et al., v. Ashley Brooke Moody et al., No. 4:21CV220-RH-MAF, 2021 U.S. Dist. LEXIS 121951 (N.D. Fla. June 30, 2021).

99. *Id*.

100. Secretary of State for Digital, Culture, Media & Sport and the Secretary of State for the Home Department, *Online Harms: White Paper,* HM GOVERNMENT (Apr. 2019), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/793360/Online_Harms_White_Paper.pdf.

101. Minister of State for Digital and Culture, *Draft Online Safety Bill*, HM GOVERNMENT (May 2021), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985033/Draft_Online_Safety_Bill_Bookmarked.pdf.

102. Department for Digital, Culture, Media & Sport, Home Office, and the Rt Hon Oliver Dowden CBE MP, *Landmark Laws to Keep Children Safe, Stop Racial Hate and Protect Democracy Online Published,* GOV.UK (May 14, 2021), https://www.gov.uk/government/news/landmark-laws-to-keep-children-safe-stop-racial-hate-and-protect-democracy-online-published.

harmful . . . ."[103] According to the U.K. Government, the Bill would also strengthen people's rights to freedom of expression.[104] Users will have access to effective routes of appeal if content is removed without good reason and companies must reinstate that content if it has been removed unfairly. Users will also be able to appeal to Ofcom. Certain popular and powerful services (Category 1 platforms) would have additional duties. They would be required to conduct and publish up-to-date assessments of their impact on freedom of expression and demonstrate that they have taken steps to mitigate any adverse effects. Platforms would be forbidden from discriminating against particular political viewpoints and will need to apply protection equally to a range of political opinions, no matter their affiliation. Journalistic content on news publishers' websites does not fall under the scope of the Bill, and articles by recognized news publishers shared on services covered by it would be exempted. Large platforms would have a "statutory duty to safeguard access to journalistic content shared on their platforms and would be held to account by Ofcom for the arbitrary removal of this content."[105]

A Polish bill published in February 2021[106] is fundamentally similar to the Florida legislation. Under this bill, social media platforms may not delete user content on the basis of their own policies, nor restrict access to user accounts. They may only use these measures if the user content violates the provisions of Polish law. In the event of content removal, or if a user account is blocked, users must be able to file a complaint on the web portal. The platform must respond to the complaint within forty-eight hours. If the complaint is not upheld, the user is also provided with a legal remedy.[107]

The European Commission submitted a legislative proposal, entitled the Digital Services Act (DSA), on December 15, 2020.[108] The proposal

---

103. *Id*.

104. *Id*.

105. John Woodhouse, *Regulating Online Harms*, PARLIAMENT.UK (Feb. 1, 2022) https://researchbriefings.files.parliament.uk/documents/CBP-8743/CBP-8743.pdf.

106. Ustawa z dnia 2021 r. o ochronie wolności słowa w internetowych serwisach społecznościowych, https://www.gov.pl/attachment/5a0c5ba6-67cb-43af-ae38-aa5dc805e14f.

107. Magdalena Gad-Nowak & Marcin S. Wnukowski, *Polish Government to Pass Law that will Allow it More Control over the Internet Content and Legitimize Blocking Access to Certain Websites*, 11 NAT. L. REV. _ (Feb. 12, 2021); Márta Benyusz & Gábor Hulkó, *Regulation of Social Media's Public Law Liability in the Visegrad States*, 1 INST. ADMIN. J. ADMIN. SCI. 6, 11–13 (2021).

108. 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. Brussels, 15.12.2020, COM (2020) 825 final, 2020/0361(COD).

does not aim to alter the liability regime of platforms, as set out in the E-commerce directive. Nevertheless, the DSA stipulates new obligations for platforms. The obligations are:

- providing information to authorities based on orders;
- designating points of contact and legal representatives;
- indicating restrictions in their terms and conditions;
- publishing annual transparency reports;
- managing notices on illegal content;
- providing reasoning for decisions;
- maintaining a complaint management system;
- ensuring the right to turn to an out-of-court body (out-of-court dispute settlement);
- processing notices on illegal content submitted by trusted flaggers with priority;
- suspending of services to recipients that frequently post manifestly illegal content;
- reporting suspicions of criminal offenses;
- publishing of more detailed transparency reports;
- user-facing transparency of online advertising.

The Digital Services Act also includes special obligations for "very large online platforms" for managing systemic risks. The proposal can be regarded as another step forward in strengthening and detailing the liability regime established by the E-Commerce Directive.

The European regulations and proposals detailed above typically remain within the framework of the notice-and-takedown procedure established by the E-Commerce Directive,[109] and are in line with its principles, only going beyond it by setting out detailed rules for it. German, Austrian, and French regulations restrict the burden on platforms to take action against illegal content, further strengthening their propensity to delete problematic user content. The envisaged U.K.

109. E-Commerce Directive, *supra* note 6.

regulation already takes aspects of freedom of speech into account, as does the E.U. proposal, which also seeks to provide for mandatory external dispute resolution, procedural guarantees, and external oversight of the platforms' activities. The Polish proposal would go even further and prohibit platforms from drawing up codes of freedom of speech and enforcing them by private means. Similar bans were contained in the defunct Florida act. There is no doubt that the freedom of platforms must also be kept in mind when creating their regulation. To date, all of the new rules or proposals succeeded in striking a (fragile) balance between restricting the intervention by platforms, protecting users' freedom and the freedom of platforms and considering the interests of other users who the platform wants to protect from offensive speech.

3. Protecting Users' Speech from Platforms: Future Considerations

The companies running social media platforms are truly the masters and governors of the communication that takes place on their platforms.[110] Of course, there are many other forums for publicity. Anyone on the Internet can express themselves free or almost free of charge, and a plethora of social media platforms are available. However, if someone wants their published opinion to have an impact on public affairs, there is not even a theoretical chance of achieving this without the big platforms. Legacy media cannot be economically viable without large social media platforms either. It has also become apparent that the big platforms are able to move together, to take steps at the same time that affect the exercise of freedom of speech, be it a general action against some harmful content, or simply banning President Trump from all platforms at the same time.[111] There is currently no viable alternative to big social media platforms in the public sphere—expression outside these platforms may almost be seen as a formal but ineffective expression of this right.

Of course, it is debateable whether it is incumbent on the state to do something about this situation, and whether it can be considered part of the constitutional protection of freedom of speech if regulation seeks to broaden the opportunities for users to actually exercise their rights. While there can be no recognized and protected right to use a given platform, the European view is that the protection of freedom of speech should not end up with public authorities unduly restricting the

110. Klonick, *supra* note 71.

111. Barrie Sander, *Democratic Disruption in the Age of Social Media: Between Marketized and Structural Conceptions of Human Rights Law*, 32 E.J.I.L. 159, 181 (2021).

freedom of their citizens. The European approach expects states to take proactive steps for the effective enforcement of the freedom of speech.[112] The democratization of the private regulation exercised by platforms, making the processes involved more transparent, creating a right of redress, and establishing external and independent oversight of the whole process by public authorities may represent the first steps in this direction.[113] The Digital Services Act proposed by the European Union has also clearly set off in this direction. The control over content on platforms should be exercised through rules that take the platform users' freedom of speech into account. In this respect, the state and legal regulation are not potential enemies, but important sponsors of freedom of speech.

Platforms may be held accountable for certain human rights standards, and aspects of content moderation in the application of private regulation may approach international legal standards that have evolved in the context of restrictions on freedom of speech.[114] This requires limitations on the decision–making freedom of the platforms as well as external oversight of the decision–making process. The first step towards this is to enforce transparency, which has been highlighted as a priority by many commentators, bearing in mind what platforms do with user content in general.[115] In the absence of adequate information, we can only gain a vague idea of the extent to which a platform interferes with the exercise of freedom of speech. At the same time, transparency is not a panacea: finding out the settings of the algorithm used and constantly modified by the platform and revealing the technical parameters to the public would be detrimental to the business interests of the platform and difficult for even experts to interpret or follow. Therefore, requiring transparency in technical terms would be largely meaningless for public authorities or individual users.

In the current circumstances, the publicity of the platforms can only be properly regulated through the cooperation of all stakeholders. The legal system typically seeks to place liability for violations in a single

112. Aleksandra Kuczerawy, *The Power of Positive Thinking: Intermediary Liability and the Effective Enjoyment of the Right to Freedom of Expression*, 8 JIPITEC 226 (2017).

113. Giovanni De Gregorio, *Democratising Online Content Moderation: A Constitutional Framework*, 36 COMPUT. L. SEC. REV. 1 (2020).

114. Thiago D. Oliva, *Content Moderation Technologies: Applying Human Rights Standards to Protect Freedom of Expression*, 20 HUM. RTS. L. REV. 607 (2020).

115. Daphne Keller & Paddy Leerssen, *Facts and Where to Find Them: Empirical Research on Internet Platforms and Content Moderation, in* SOCIAL MEDIA AND DEMOCRACY (Nathaniel Persily & Joshua A. Tucker eds., 2020); Robert Gorwa & Timothy Garton Ash, *Democratic Transparency in the Platform Society, in* SOCIAL MEDIA AND DEMOCRACY (Nathaniel Persily & Joshua A. Tucker eds., 2020).

location by clearly designating the addressee of a rule. This approach is already being sidelined in E.U. legislation, for example, which makes it mandatory for platforms to delete unlawful content if certain conditions are met. Public authorities on the one hand determine what qualifies as unlawful and impose direct liability on the infringing user, and on the other hand monitor compliance with the obligation imposed on the platforms. In principle, without going into detail, this system may seem necessarily respectful of freedom of speech and democratic publicity, but in practice it typically results in platforms deleting content that has been flagged as unlawful, thus avoiding the hassle of thorough investigation. The actual significance of this procedure is amplified because the regulation does not prohibit private regulation by platforms, so they may delete content that they consider undesirable to a much greater extent. Content that is considered unlawful by legal regulation in most cases also qualifies as deletable according to the private regulation of the platform—with certain exceptions. For example, Facebook does not ban defamatory content, while the law does, so the notice-and-takedown procedure is often irrelevant. However, it seems certain that the future of regulation will see the division of legal responsibility between stakeholders by keeping in mind the cooperative responsibility.[116] Users must be directly responsible for the content they make accessible, and the platform must respect both the user's freedom of speech and the interests of the democratic public. Platforms must accept that media-like businesses that are assigned social responsibility need to strengthen the freedom of discourse on the Internet.[117]

In recent years there have been several regulatory attempts to provide answers to problems raised in connection with the operation of platforms. The German, Austrian, or French forms of regulation introduced above increase the existing burden on platforms, reinforcing the need for management of content and its deletion where necessary. This may seem to be an adequate response to the social tensions experienced in those states, but it has the effect of further restricting the enforcement of freedom of speech while indirectly strengthening the legitimacy of the private regulation of platforms. In the meantime, Facebook has already pre-emptively set up a body that it presumably hopes will dampen governments' regulatory zeal. Facebook's recently created self-regulatory mechanism oversees its own operation by setting up a supervisory body independent of the platform, the state, and other

116. Natali Helberger et. al., *Governing Online Platforms: From Contested to Cooperative Responsibility*, 34 INFO. SOC'Y. 1, 1–8 (2018).

117. Balkin, *supra* note 69, at 1209.

industry players alike.[118] This Oversight Board, previously referred to by Mark Zuckerberg as the "Supreme Court of Facebook," is not intended to serve as an appeal forum for individual cases but as a body that sets general benchmarks for freedom of expression.[119] An essential element of self-regulation is the separable nature of the regulated and the regulator: the Oversight Board (the regulator) will therefore be considered self-regulatory if Facebook (the regulated) actually submits itself to its decisions. On the other hand, the Oversight Board may be considered to be private regulation, as its activities affect the freedom of expression of the platform users. The rules of operation of the Board are established by Facebook; its members are appointed by Facebook; and its competence extends exclusively to the Facebook platform, which all undermine its independence. The establishment of the Oversight Board is another step towards the strengthening of the private regulation that has been developing in parallel with the legal system.

Settling the relationship between legal and private regulation would require a clear distinction between unlawful content and that which is merely harmful. While all unlawful content is harmful, provided that the ban on it complies with democratic standards, harmful, damaging content is not necessarily unlawful. At present, regulation in Europe tolerates, and in some cases even supports, platforms taking action against not-unlawful but harmful content, for example, in the case of action against fake news. In the meantime, it has not attempted to set clear criteria to prevent interference with the exercise of freedom of speech from being arbitrary.[120] The nature of harmful content is not determined by democratic procedures. Instead, the platforms decide what they consider to be harmful.[121] The Florida experiment presented above and the Polish bill represent a step in the direction of pushing back against private regulation. The intention behind such initiatives is to be welcomed, even if the experiments themselves are far from perfect. The Polish bill aims to make private codes compiled by platforms irrelevant, with platforms only authorized to remove content that is prohibited by the legal system, namely unlawful content. This would expect the platforms' moderators and algorithms to become familiar with the country's freedom of speech legislation, which is

118. Makena Kelly, *Facebook's Oversight Board will Include a Former Prime Minister and Nobel Prize Winner*, THE VERGE (May 6, 2020), https://www.theverge.com/2020/5/6/21249427/facebook-oversight-board-nobel-peace-prize-instagram-snowden.

119. Kate Klonick, *The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression*, 129 YALE L. J. 2232 (2020).

120. Sander, *supra* note 111, at 179.

121. Helberger, *supra* note 116, at 8.

undoubtedly more difficult than enforcing worldwide private regulation that uses rules created by the platform. The Polish solution may not be operable (no guarantees would exist even at the European level, given the varying restrictions on freedom of speech) although its underlying principle may be attractive to friends of freedom of speech.

C. The Possible Application of the Public Forum Doctrine

The well-established public forum doctrine has enshrined the right to use certain physical spaces used by the community for the purposes of exercising freedom of speech. That right presupposes the existence of community fora where that freedom may be exercised. According to the case law of the Supreme Court of the United States, freedom of speech can, to some extent, be exercised in the new public fora. However, this practice is not fully uniform and recognizes the right to freedom of speech in the new public fora only subject to strict restrictions. In *Marsh v. Alabama*,[122] a factory banned the distribution of flyers in its own "town," built to accommodate its own workers, and the Court found this restriction unconstitutional. Based on the findings of the judgment delivered in *Pruneyard Shopping Center v. Robins*,[123] freedom of speech may also be exercised in a private shopping mall with certain restrictions, as long as it made clear that the opinion expressed is not that of the owner of that institution. This right obviously does not include the organization of gatherings in the building, for example. In *International Society for Krishna Consciousness v. Lee*[124] the Court ruled that although a privately-owned airport is not a public forum, distribution of flyers is allowed on its territory. However, addressing people with a view to raising donations is not.

European legal systems also guarantee the freedom of streets and spaces for the purpose of assembly and free speech, but in general privately owned buildings and real estate cannot be taken over for the purpose of exercising freedom of speech. In *Appleby v. the United Kingdom*,[125] the applicant stated that the owner of a private shopping mall had violated his rights by not consenting to the installation of a table and a podium in the building to collect signatures and distribute leaflets (the protest would have targeted construction works in a town park). The European Court of Human Rights held there was no violation of freedom of speech and right of assembly on the side of the state. Even so, large shopping malls can hardly be regarded as entirely

122. 326 U.S. 501 (1946).

123. 447 U.S. 74 (1980).

124. 505 U.S. 672 (1992).

125. Eur. Ct. H.R. Application no. 44306/98, judgment of 6 May 2003.

private institutions with regard to the exercise of freedom of speech, so it may be justified to ensure this right can be exercised on their premises to some extent. The general European interpretation of freedom of assembly and speech holds that only state-owned and local government-owned areas may be used to exercise these rights. The total exclusion of privately owned properties from this circle limits a significant amount of speech and its effectiveness, as well as being potentially discriminatory since the owner can freely decide who can enter the area and who is excluded from it. In this way, the protection of private interests may be contrary to the public interest.[126]

Social media platforms do not exist in a physical space and the service is privately owned by the company providing it, yet they are used by millions of people on a daily basis for the purpose of exchanging information and expressing opinions. Social media platforms may also be used to commit criminal offenses. In this context, it is an interesting question whether certain users may be banned from such platforms on the basis of crime prevention considerations. Sex offenders constitute a group of particular importance, as they can easily contact minors via the platforms. John Hitz argues that a general ban on those convicted of such offences (following the enforcement of their punishment) would be inconsistent with freedom of speech.[127] Once rehabilitated, sex offenders may exercise their freedom of speech without any restriction.[128] According to U.S. doctrine, banning such persons from commonly used platforms (not only those exclusively used by minors) would constitute a content-neutral restriction on speech that is not tailored narrowly enough.[129]

This very issue was considered by the Supreme Court of the United States in *Packingham v. North Carolina*,[130] which is the second most important decision on the freedom of online speech since *Reno v. American Civil Liberties Union*[131] twenty years earlier. Justice Kennedy, drafter of the majority opinion, described the Internet as the "modern public square," where members of the public exchange opinions.[132] While the lower court considered the law of North Carolina

126. Jacob Rowbottom, *Property and Participation: A Right of Access for Expressive Activities*, 2 EHRLR 186 (2005).

127. John Hitz, *Removing Disfavored Faces from Facebook: The Freedom of Speech Implications of Banning Sex Offenders from Social Media*, 89 IND. L. J. 1327 (2014).

128. *Id*. at 1341.

129. *Id*. at 1349–56.

130. 137 S. Ct. 1730 (2017).

131. 521 U.S. 844 (1997).

132. *Packingham*, 137 S. Ct. at 1737.

as a restriction on action but not on speech, and consequently applied a less stringent standard. The Court disagreed and changed the ruling, as its restriction on speech was not of a narrowly tailored character.[133]

By comparing the Internet to a physical space, Justice Kennedy raised the issue of whether the public forum doctrine could be applied with regard to the Internet. U.S. law distinguishes three different kinds of public forums, each of which is subject to different standards concerning the limitation of speech. The first kind includes traditional public forums (public squares, streets, and parks), where the exercise of the freedom of speech is customary. The second kind includes designated or limited public forums, which are traditional but specifically designated places for exercising the freedom of speech (for example, conference halls that can be used with the permission of their owner or manager). The third kind of public forum includes non-public forums that do not serve as a place where anyone can speak but where speech takes place nonetheless (such as hospitals, prisons, and military bases).[134] More and more restrictions on free speech may be applied to the different kinds of public forums, proceeding from the first to the third category.

In his concurring opinion to the *Packingham* judgment, Justice Alito suggested that the Court might be wrong to compare the Internet per se to streets and public parks.[135] The comparison may be valid regarding social media platforms used by government organs and bodies, but most of the communication and exchange of ideas conducted through such platforms is private in nature,[136] meaning that the doctrine of public forums cannot be applied, and the platforms in general, as well as their individual users (regarding their own profile), are free to adopt their own rules of speech. According to this approach, user access to such platforms may not be prohibited by the government using legal means. However, the service provider may certainly do so without any limitation, with possible exceptions when applying anti-discrimination rules. This approach may be challenged in that it allows a platform to act as a kind of government in itself, meaning that the freedom of speech should also be guaranteed with regard to the restrictive practices of the platform.[137] However, such requirements would come close to challenging the ownership rights of the platforms themselves.

---

133. *Id*. at 1736–38.

134. Lyrissa Lidsky, *Public Forum 2.0*, 91 B.U. L. REV. 1975, 1980–92 (2011).

135. *Packingham*, 137 S. Ct. at 1743.

136. *First Amendment-Freedom of Speech-Public Forum Doctrine-Packingham v. North Carolina*, 131 HARV. L. REV. 233, 238 (2017).

137. Klonick, *supra* note 71, at 1609.

According to the traditional approach to freedom of speech, a social media platform is similar to a privately-owned shopping center, in that its owner or operator can ban or remove any person from the premises if the rules of using the property as determined by the owner have been violated.[138] This might result in a situation where services that once promised to facilitate the exercise of individual freedoms enter into deals with oppressive regimes in the hope of business advantages.[139]

Social media platforms are also used by public institutions and officials to provide information, collect opinions, and so forth, and the public profile of a politician or a local government may be subject to different rules than those of private users. Former President Donald Trump was a prominent user of Twitter, where he used to block those users who posted critical comments under his tweets. The question is whether he (or the actual manager of his account) could prohibit others from reading his messages (that is, "following" him), as numerous U.S. citizens have experienced. According to the dominant approach, the public forum doctrine may not be applied concerning the relationship between Twitter and its users, but it may be applied concerning the relationship between two users, in particular if one of them is an elected public official. In other words, an area of a social media platform may be considered a public forum if it is used for public political communication.[140]

This approach was illustrated by a district court in *Knight First Amendment Institute v. Trump*.[141] Numerous plaintiffs filed a lawsuit against the former President and the White House staffer managing his Twitter account because they had been banned from following that account. As a result of the ban, they could not have direct access to and comment on the former President's tweets. Nor could they read the related comments. They could only learn about communications made by the former President through comments made by their contacts. After analyzing the applicability of the public forum doctrine, the court ruled that the former President's account was a designated or limited forum from which a person whose speech did not cross the limits of the freedom of speech could not be banned (the plaintiffs were banned because of their tweets that disputed the content of the presidential

138. *Facebook is not the Public Square*, THE NEW YORK TIMES (Dec. 25, 2014), https://www.nytimes.com/2014/12/26/opinion/facebook-is-not-the-public-square.html.

139. Mike Isaac, *Facebook Said to Create Censorship Tool to Get Back into China,* THE NEW YORK TIMES (Nov. 22, 2016), https://www.nytimes.com/2016/11/22/technology/facebook-censorship-tool-china.html.

140. Lidsky, *supra* note 134, at 1994–2002.

141. 302 F. Supp. 3d 541 (S.D.N.Y. 2018).

tweets).[142] The White House is not the owner of the service, not even in the context of the single Trump account concerned, but the account is operated under its control and supervision, which is enough to consider it a public forum.[143] The restriction of political opinions is unacceptable in such a forum[144] (meaning, *a contrario*, that even a political figure might be banned from the service if he or she crosses the limits of free speech).

The second-instance judgment in this case[145] confirmed the first-instance decision that the former President could not block users who criticized or mocked his policy from following his own Twitter account. The President's Twitter account is a public forum, and its manager may not restrict access to those who post opinions on it that are not otherwise unlawful or violate the provisions binding the users of the platform.

The Supreme Court of the United States vacated this ruling, as Trump was no longer president, and in January 2021 Twitter permanently suspended his account.[146] Although the decision was expected, the concurring opinion of Justice Thomas raises important points.

> Respondents have a point, for example, that some aspects of Mr. Trump's account resemble a constitutionally protected public forum. But it seems rather odd to say that something is a government forum when a private company has unrestricted authority to do away with it. The disparity between Twitter's control and Mr. Trump's control is stark, to say the least. Mr. Trump blocked several people from interacting with his messages. Twitter barred Mr. Trump not only from interacting with a few users, but removed him from the entire platform, thus barring all Twitter users from interacting with his messages. [147]

Justice Thomas's suggestion is valid and, in the light of subsequent events, as well as President Trump's ban from the platform,[148] it is

142. *Id.* at 572–73.

143. *Id.* at 565–69.

144. *Id.* at 575–76.

145. Knight First Amendment Inst. at Columbia Univ. v. Trump, 928 F.3d 226 (2nd Cir. 2019).

146. Biden v. Knight First Amendment Inst. at Columbia Univ., 141 S. Ct. 1220 (2021).

147. *Id.* at 2.

148. Kate Conger and Mike Isaac, *Twitter Permanently Bans Trump, Capping Online Revolt*, THE NEW YORK TIMES, (Jan. 8, 2021). https://www.nytimes.com/2021/01/08/technology/twitter-trump-suspended.html.

difficult to argue that the account of a political actor constitutes a public forum if the platform was able to treat it without restrictions or even close it down. This possibility, of course, existed even before the President was banned; in that sense, Justice Thomas's suggestion was not linked to specific events. Moreover, instead of considering the platforms as public forums, Justice Thomas raised the applicability of the category of "places of public accommodation." In other words, the idea that platforms are public places which provide publicly accessible services. The doctrine of these services may restrict the platform's right to terminate the provision of the service unilaterally.[149]

An appellate court ruled similarly to the *Knight* judgment in *Davison v. Randall*,[150] noting that the official Facebook page operated by the defendant, the chairman of a county school council, is a public forum. The deletion of the plaintiff's critical comments was an impermissible restriction of his freedom of speech, which constituted discrimination on the basis of point of view, and thus the defendant had violated a constitutional right. However, the platform as a whole is not a public forum, and its provider is not a public actor which must ensure equal access to the platform it owns. The deletion of the Facebook account of Russian trolls by the service provider does not therefore constitute a restriction of freedom of speech. Section 230 of the CDA or its contract with users does not restrict the platform provider from taking action against content that is incompatible with its policy but not necessarily unlawful.[151]

If large social media platforms are considered public fora from a legal point of view, it would open the gates to wider restrictions on their operation. Knowing the role they play in publicity and their *de facto* unavoidable nature, this does not seem unthinkable.[152] This would allow private regulation by platforms to be prohibited, while the restrictions on freedom of speech would remain applicable. However, declaring them to be public fora would require a significant reinterpretation of the doctrine, and interference with the exercise of the right to private property to an extent that would hardly be

149. *Id.*

150. 912 F.3d 666 (4th Cir. 2019).

151. Federal Agency of News LLC, et al. v. Facebook, Inc., 432 F. Supp. 3d 1107 (N.D. Cal. 2020).

152. Trevor Puetz, *Facebook: The New Town Square,* 44 S.W. L. REV. 385 (2014).

compatible with the currently prevailing conception of the First Amendment.[153]

D. The Horizontal Effect of Fundamental Rights—The German Example

The doctrine of the horizontal effect of fundamental rights is capable of protecting the user vis-à-vis the platform, in their contractual relationship. The roots of this doctrine are in Europe.

> The [German] doctrine of third party effect (*Drittwirkung*) instantiates the idea that the (economic) constitution entails legal obligations on private law interactions of private persons in their relationships *inter se* . . . . *Drittwirkung* (third party effect) may be direct or indirect . . . . Horizontal direct effect is the application of public law rules to directly affect legal relations between private individuals in their relations with other private law persons.[154]

The U.S. legal system calls this "state action doctrine" rather than "horizontal effect," where the law obligates private parties to respect the constitutional rights of others.[155] With regard to online platforms, intervention in the relationship between a platform and a user would not fit into the U.S. legal system.[156]

A private law approach to platforms and their users may be more expedient than contemplating strict state regulation of platforms. Applying the doctrine of the horizontal effect of fundamental rights could help to restore the imbalance between the platform and the user, forcing platforms to take their users' right to freedom of speech into account in their decisions. This does not require legal authorization either. Courts may apply this approach in legal disputes, provided that this is not contrary to the traditions, practices, and principles of the legal system they operate in.

Two recent German cases clearly show the contradictions and ambiguity which arise when applying the constitutional free speech doctrines to a contractual relationship between a social media platform

153. Matthew P. Hooker, *Censorship, Free Speech & Facebook: Applying the First Amendment to Social Media Platforms via the Public Function Exception*, 15 WASH. J. L. TECH. & ARTS 36 (2019).

154. Eric Engle, *Third-Party Effect of Fundamental Rights (Drittwirkung)*, 5 HANSE L. REV. 165 (2009); *see* Mark Tushnet, *The Issue of State Action/Horizontal Effect in Comparative Constitutional Law*, 1 INT'L. J. CONST. L. 79 (2003).

155. Richard S. Kay, *The State Action Doctrine, the Public/Private Distinction, and the Independence of Constitutional Law*, 10 CONST. COMMENT. 329 (1993).

156. Jonathan Peters, *The "Sovereigns of Cyberspace" AND State Action: The First Amendment's Application—or Lack Thereof—to Third-Party Platforms*, 32 BERKLEY TECH L.J. 989 (2017).

and its user. In *Themel v. Facebook Ireland, Inc.,*[157] the Higher Regional Court in Munich, Germany, held that the deletion of the plaintiff's comment by Facebook constituted a breach of contract, as the platform was required to respect her right to freedom of expression under Article 5 of the German Constitution (Grundgesetz).[158] The facts of the case were that on August 7, 2018, *Spiegel-Online*, a German news website, posted an article on its Facebook page entitled "Austria announces border controls." There was a harsh debate in the comments under the Facebook post, and Heike Themel, a German politician and member of the right-wing AfD Party, was referred to as a "Nazi-slut." She responded to that comment, quoting a German poem: "I can't compete in an argument with you. You are unarmed and this wouldn't be very fair from my side." Facebook deleted the comment and suspended her account for thirty days. This decision was based on Facebook's Community Standards, rule 5.2, which prohibits hate speech on the platform. After she approached the court, it held that the application of rule 5.2 violated Section 241(2) of the German Civil Code,[159] which says that "[a contractual] obligation may also, depending on its contents, oblige each party to take account of the rights, legal interests and other interests of the other party."[160] As the Community Standards give Facebook the power to decide on its own which posts or comments violate its rules, the court noted that this power contradicts the Civil Code's requirement. The court emphasised that Facebook as a social media platform provides a "public marketplace" for an exchange of views and opinions, and that legally permissible expressions cannot be deleted from the platform. As Heike Themel's comment did not constitute hate speech, Facebook's deletion of the comment and suspension of Themel's account was unlawful.

In another German case, *User v. Facebook Ireland, Inc.*,[161] the Regional Court in Heidelberg, Germany, arrived at the completely opposite conclusion. The court rejected the plaintiff's argument that her right to freedom of expression had been infringed. Prior to the decision, in July 2018, a Facebook user commented below a post concerning integration of migrants in Germany: "[r]espect! That is the keyword! Fundamentalist Muslims regard us as soft grown heathens,

157. Themel v. Facebook Ireland, Inc., 18 W. 1294/18, 24 August 2018, Higher Regional Court Munich, Germany.

158. German Constitution, Art. 5 (2012).

159. German Civil Code BGB § 241(2).

160. *Id.*

161. User v. Facebook Ireland, Inc., 1O 71/18, 28 August 2018, Regional Court in Heidelberg, Germany.

pig-gluttons and our women as whores. They do not respect us." On July 16, 2018, Facebook deleted the user's comment and blocked her profile for 30 days. After Facebook refused to reverse its decision, the plaintiff sought a preliminary injunction before the Regional Court in Heidelberg. The central issues before the court were whether Facebook was entitled to remove the post and block the user, and whether Facebook's Community Standards were consistent with Section 307 of the Civil Code, which says in its Paragraph 1 that provisions in standard business terms are ineffective if, contrary to the requirement of good faith, they unreasonably disadvantage the other party to the contract with the user. An unreasonable disadvantage may also arise from the provision not being clear and comprehensible.

The court noted that Facebook's standards list the types of expression that are not protected and define the boundaries of what can be considered restricted speech. In addition, the rules indicate the kind of consequences each user faces if he or she violates these standards. Accordingly, the court held that the standards cannot be considered non-transparent, and they did not discriminate against users inappropriately. As a conclusion, the court found that Facebook's rules adequately take into consideration the right to freedom of expression and that, even though aggressive opinions or extreme expressions are protected under the Constitution, Facebook as a private party does not have to grant its users the full right to freedom of expression that is provided by the state in the constitutional context.[162]

These two decisions take two different paths. The latter decision fits into the usually applied legal framework, which—through the recognition of the platform's property and free speech rights—allows Facebook to delete more or less any users' content it finds inappropriate. The former one aims to restrict the platform's powers in this regard. The decisions depict the possible strengths and weaknesses of mandating the law of contracts to resolve free speech issues arising between private parties.

In other cases, German courts have also ruled that platforms are not allowed to delete lawful and protected opinions, in view of the horizontal effect of fundamental rights.[163] Under the German constitution, the state has a duty to ensure the freedom of speech for its citizens, even at the expense of private parties (platforms). Although

---

162. *See* Colombia Global Freedom of Expression, *User v. Facebook Ireland, Inc.*, COLUMBIA UNIVERSITY (Aug. 28, 2018), https://globalfreedomofexpression.columbia.edu/cases/user-v-facebook-ireland-inc.

163. Matthias C. Kettemann & Anna S. Tiedeke, *Back Up: Can Users Sue Platforms to Reinstate Deleted Content?*, 9 INTERNT POL'Y. REV. 1, 10 (2020).

the contract between the platform and the user may allow the former to restrict the freedom of speech, the German courts consider that these provisions are void if they put the user into a significantly disadvantaged position by violating the principle of good faith in private law.[164]

The German Constitutional Court has also heard a case wherein the applicant asked the court to restore a comment he posted that had been deleted by the platform as hate speech and to reactivate his account, which had been suspended at the same time. The court started from the provision of the Basic Law according to which "[a]ll persons shall be equal before the law." [165] According to the court's decision: "[u]nder specific circumstances, Article 3(1) of the Basic Law may give rise to requirements pertaining to the right to equality in the context of relationships between private actors."[166] The court in its reasoning performed a weighing of the disadvantages for the involved parties, and found that the consequences that would occur if the interim injunction was not issued but the main proceedings—on whether the given content can be deleted or not by the platform—were successful would outweigh the disadvantages that would arise if the interim injunction was issued, but the main proceedings proved to be unfounded. Therefore, it granted the preliminary injunction. In this way, the court ordered a temporary delay, but set an example of how—besides private law—constitutional equal rights protection may also be used to limit the power of online platforms.

Given the key role of social media platforms in public communication, it would not be completely alien to the European approach to view them as service providers, which would open the door to the recognition of horizontal effect.[167] In any case, the draft E.U. Digital Services Act also contains a number of provisions that limit the possible content of the contract between the platform and the user, with a view to protect the user's interests. Such a requirement would involve an obligation for the platform to justify its decisions made on user content, to operate a complaint handling mechanism, and for the contract to provide for the right to independent out-of-court dispute resolution.

---

164. *Id*. at 10–11; *see* German Civil Code, [BGB] § 307.

165. Second Chamber of the First Senate, Order of 22 May 2019, 1 BvQ 42/19.

166. *Id.*

167. Giovanni De Gregorio, *From Constitutional Freedoms to the Power of the Platforms: Protecting Fundamental Rights Online in the Algorithmic Society*, 11 EUR. J. LEGAL STUD. 65, 101 (2019).

Nicolas Suzor calls efforts to protect users' rights "digital constitutionality,"[168] and he approaches it not from the perspective of private law and contract law but suggests the application of certain principles of the "rule of law" to the relationship between the platform and the user, such as by applying the criteria of it being consensual, transparent, equally applied, and relatively stable. The aim of digital constitutionality is to revisit how private governance can be limited,[169] and the result may be similar to if it were approached from a private law perspective: a restriction on the platform's freedom of choice, aiming to protect the user's freedom of speech.

IV. POTENTIAL OBLIGATIONS OF SOCIAL MEDIA PLATFORMS TO PROTECT THE PUBLIC INTEREST

The European notion of freedom of the press maintains that the media have "duties and responsibilities" in the course of exercising this freedom. This wording is included in two documents that play an important role at the international level in the interpretation of freedom of speech and the press: The European Convention on Human Rights (Article 10(2)) and the International Covenant on Civil and Political Rights (Article 19(3)), which also applies outside the community of European states. Nevertheless, the European court of Human Rights (ECtHR), which enforces compliance with the European Convention on Human Rights, clearly interprets these duties in relation to the press,[170] and it does not emphasize them in the case of individual speakers.[171] These duties and responsibilities are applied towards democratic publicity, and it is the media's duty to report on important issues of public interest and to respect the rights of others. However, the former is mostly mere rhetoric, as the media in Europe do not have a general duty to serve the public interest and are free to report on what they want, or to ignore any event or opinion, so it is possible to establish a newspaper, to operate a television channel or to communicate perfectly non-political tabloid news on any topic. However, with regard to media services such as television, radio, and, to a lesser extent, on-demand media services, European (national)

168. Nicolas Suzor, *Digital Constitutionalism: Using the Rule of Law to Evaluate the Legitimacy of Governance by Platforms*, 4 SOC. MEDIA + SOC'Y. 1 (2018).

169. *Id*. at 4.

170. Eur. Ct. H.R. Sunday Times v. the United Kingdom (No. 2), No. 13166/87, judgment of 26 November 1991; The Observer and Guardian v. The United Kingdom, No. 13585/88, judgment of 26 November 1991; Jersild v. Denmark, No. 15890/89, judgment of September 23, 1994 [GC].

171. JAN OSTER, MEDIA FREEDOM AS A FUNDAMENTAL RIGHT 36 (2015).

regulations may, if certain conditions are met, impose obligations to provide citizens with adequate information. These rules are mostly absent from the U.S. legal system. Nevertheless, the application of the principles that underlie the rules intended to protect the public interest to social media platforms may be worth reconsidering.

A. The Social Media as "Media"

Large Internet gatekeepers usually consider themselves tech companies.[172] It is in their best interests to do so for two reasons. First, the regulations applicable to technology companies are far narrower and less stringent than those applicable to media companies (which are also subject to content regulation, special restrictions on competition, the prohibition of concentration, and the obligation to perform public interest tasks). Second, the moral requirement of social responsibility is far less frequently mentioned concerning the activities of tech companies. However, the legal classification of a given service does not derive from the self-image of the service provider but from the nature of its activities. For this reason, some of the gatekeepers, primarily the social media platforms, can be considered media undertakings.[173]

Previously, Facebook insisted that its service is nothing but a neutral platform, and the company does not have anything to do with how or for what purpose it is used by users.[174] Discussions conducted through the platform may improve participation in elections, but the service itself does not influence the outcome of elections in any way.[175] It is more like a billboard: anybody can sign or display anything on it. In light of the events that unfolded in recent years, this position does not seem easy to maintain any longer. It was revealed in the Spring of 2016 that Facebook's Trending Topics service distorts the significance of certain pieces of news on a political basis, so that some content was presented as if it were more or less important than it actually was. The distortion

172. Ashley Rodriguez, *Zuckerberg Says Facebook will Never be a Media Company—Despite Controlling the World's Media*, QUARTZ (Sept. 1, 2016), https://qz.com/770743/zuckerberg-says-facebook-will-never-be-a-media-company-despite-controlling-the-worlds-media.

173. Charles Warner, *Fake News: Facebook is a Technology Company*, FORBES (Nov. 27, 2016), https://www.forbes.com/sites/charleswarner/2016/11/27/fake-news-facebook-is-a-technology-company/#7b65816e1381.

174. Nicholas Thompson & Fred Vogelstein, *Inside the Two Years That Shook Facebook—And the World*, WIRED (Feb. 12, 2018), https://www.wired.com/story/inside-facebook-mark-zuckerberg-2-years-of-hell.

175. Alexis C. Madrigal, *The False Dream of a Neutral Facebook*, THE ATLANTIC (Sept. 28, 2017), https://www.theatlantic.com/technology/archive/2017/09/the-false-dream-of-a-neutral-facebook/541404.

was deliberate and implemented manually, that is, not caused by a possibly miscalibrated algorithm.[176] After the U.S. presidential election of 2016, the platform was widely accused of not having done anything to prevent the spread of false news, thereby contributing to the victory of Trump and the defeat of Hillary Clinton.[177] Around the same time, a "fake news factory" was discovered in Macedonia,[178] followed by news of meddling in social media by Russian intelligence services, sparking endless investigations.[179]

Compounded by the scandal concerning Cambridge Analytica in 2018 (which also touched upon the debate on the protection of users' personal data),[180] Facebook could not deflect these accusations by disclaiming any responsibility for its users, and it could not claim to remain neutral any longer. By 2018, the platform realized that similar to the media and other publishers, it bears responsibility toward the public for the state of democracy.[181] It is another issue how a series of such problems could be handled (if they can be handled at all), considering that the platform was designed to rapidly provide a wide audience to all statements, including false ones. It should be noted here that Facebook was not a neutral platform even before 2016. Its algorithms produce a personalized news feed for each and every user, using settings that are dependent on but not entirely under the control of the user concerned. Naturally, the news feed settings serve the business interests of the platform, which are legitimate interests but unlikely to foster any neutral behaviour.

Philip Napoli and Robyn Caplan offer a summary of the questions that arise in this field. The authors argue that, considering their main activities, large online gatekeepers should no longer be considered tech

176. Sam Thielman, *Facebook News Selection is in Hands of Editors Not Algorithms, Documents Show*, THE GUARDIAN (May 12, 2016) https://www.theguardian.com/technology/2016/may/12/facebook-trending-news-leaked-documents-editor-guidelines.

177. Antonio G. Martínez, *How Trump Conquered Facebook—Without Russian Ads*, WIRED (Feb. 23, 2018), https://www.wired.com/story/how-trump-conquered-facebookwithout-russian-ads.

178. Samanth Subramanian, *Welcome to Veles, Macedonia, Fake News Factory to the World*, WIRED (Feb. 15, 2017), https://www.wired.com/2017/02/veles-macedonia-fake-news.

179. Sheera Frenkel & Katie Benner, *To Stir Discord in 2016, Russians Turned Most Often to Facebook*, THE NEW YORK TIMES (Feb. 17, 2018), https://www.nytimes.com/2018/02/17/technology/indictment-russian-tech-facebook.html.

180. *See The Cambridge Analytica Files*, THE GUARDIAN, https://www.theguardian.com/news/series/cambridge-analytica-files.

181. Thompson & Vogelstein, *supra* note 175.

companies.[182] The identity of these companies is based on the argument that they do not produce any content themselves but merely facilitate the publication of content created by their users (social media), or provide links to such content upon request by a user (search engines). This may be true, but Napoli and Caplan identified a number of features that make the operation of such companies quite similar to that of the media. From the perspective of the public sphere, the distribution of content is also of great importance, in addition to that of content creation, and such activities used to form part of the activities of media companies before the emergence of the Internet. These companies employ human workers, either to make certain editorial decisions or to configure the algorithms that make such decisions automatically, and such editorial decision-making is an essential part of their services. Similar to legacy media, the services of these companies seek to provide the members of their audience (their users) with whatever they want to see, meaning that they may not be considered neutral platforms. Last but not least, the main source of income of these companies is advertising—just like the media.[183] The operators of the most influential online platforms—such as Facebook and Google—are considered media companies by legal scholars (even if not by existing legal doctrine).[184]

B. The Free Speech of the Platforms

It seems clear now that platform providers are similar to traditional media in terms of making editorial decisions, and therefore they have a right to free speech by way of selecting and sorting pieces of content (including giving their users power to influence these decisions according to their own preferences). In a sense, users' individual news feeds, as edited by the platforms' algorithm, are Facebook's opinions on what its users might be most interested in and how the platform's business interests could be best served in that context. "With the curated production of news stories, editorial control is no doubt exercised, and thus the controller of the algorithm would be a publisher."[185] If a platform has an opinion, it is afforded protection

182. Philip M Napoli & Robyn Caplan, *Why Media Companies Insist They're Not Media Companies, Why They're Wrong, and Why it Matters*, FIRST MONDAY (Apr. 13, 2017), http://firstmonday.org/ojs/index.php/fm/article/view/7051/6124.

183. *Id*.

184. *See*, *e.g.*, Jack M Balkin, *Old-School/New-School Speech Regulations*, 127 HARV. L. REV. 2296, 2304 (2014).

185. *Id*. at 1373. On content selection by algorithms as speech, *see* Benjamin, *supra* note 30; Benjamin, *supra* note 37.

under the constitutional rules, but it may also be subject to restriction, pursuant to applicable legal principles.

As noted throughout this paper, the activities social media platforms carry out concerning the user content they manage is not neutral. Either because they seek to comply with legal regulations or since they act on their own initiative, the operators of these platforms carry out a kind of editorial task that involves the assessment and evaluation of such content. The private regulation implemented by platforms is not value-neutral either, as it clearly reflects their objective to accommodate as many users as possible. This objective is not always compatible with the goal of acting as a robust defender of free speech. Moreover, through this private regulation and the necessary prioritization of various user content, platforms exercise real opinion-forming power, the dimensions of which have never before been experienced in the public sphere.[186]

Social media platforms usually seem to be the champions of free speech. But, as Bernal notes, "in practice free speech is just a tool for them. They will champion it when it suits them and not champion it when it does not."[187] Facebook is often compared to a nation state.[188] Even though social media platforms have far more extensive ways of modeling private regulation than a nation state, the analogy applies in that nation states are neutral from a religious or philosophical point of view, but they are not value-neutral. A social media platform can also make value-based decisions and could ban hateful people from its system. If we accept the public sphere to be a fundamental institution (it would be difficult not to do so), surrendering ideological neutrality and embracing bias would lead to serious problems, even though it cannot be prohibited using the currently available legal and regulatory means. However, Western countries are democracies, meaning that the limits of free speech—among other things—are set out and made in compliance with the rules as supervised by elected officials, courts, and other authorities operating within a framework of constitutional safeguards and guarantees. If a social media platform were a state, it most certainly would not be a democracy.

According to the (hardly surprising) findings of a survey, the owners and executives of U.S. tech companies established in Silicon Valley hold liberal, cosmopolitan, and globalist political views and support the

186. Natali Helberger, *The Political Power of Platforms: How Current Attempts to Regulate Misinformation Amplify Opinion Power*, 8 DIGIT. JOURNALISM 842 (2020).

187. PAUL BERNAL, THE INTERNET, WARTS AND ALL: FREE SPEECH, PRIVACY, AND TRUTH 127 (2018).

188. Anupam Chander, *Facebookistan*, 90 N.C. LAW. REV. 1807 (2012).

extension of human rights—in all matters that do not interfere with their business interests. However, they are against government regulation in general, and labor and employment policies in particular, on which matters they tend to agree with conservative libertarians.[189] With regard to Facebook, signs of ideological bias also exist. The manifesto published by Mark Zuckerberg in 2017 clearly reflects a political agenda (albeit a rather naive one, considering the chances of its implementation) to build a global community above and beyond nation states.[190] The wording goes beyond the goal of providing a safe space and envisages the abandonment of the concept of nation states. Naturally, this idea is not new in the era of globalization, but, reading between the lines, one might find the objective of surpassing "national" societies an aim that is still somewhat surprising in the age of world trade, international organizations, and an increasingly united Europe.

It became clear during several scandals in the previous years (see later) that Facebook is capable of exerting direct political influence, even without any noble cause or a publicly acknowledged ideological stance. The owners and executives of social media platforms can exercise their freedom of speech, including making decisions concerning the infrastructure they own. But a platform can also be harmful to democratic public life if it grows really large but fails to manage debates conducted on the platform with due regard to the notion of the marketplace of ideas, that is if it attempts to influence such exchanges using obscure means that lack transparency. However, even those arguing that the selection of user content by platform algorithms constitutes speech in itself acknowledge that the regulation of platform algorithms for this selection is not prohibited by the constitutional rules of freedom of speech.[191]

C. Platform Regulation in the Interest of the Public

1. The Problem of Content Diversity

The removal of undesirable content is not the only means of implementing private regulation. A far more powerful means is the editing and sorting of content presented to individual users, as well as

189. Farhad Manjoo, *Silicon Valley's Politics: Liberal, With One Big Exception*, THE NEW YORK TIMES, (Sept. 6, 2017), https://www.nytimes.com/2017/09/06/technology/silicon-valley-politics.html.

190. Mark Zuckerberg, *Building a Global Community*, FACEBOOK, https://www.facebook.com/notes/mark-zuckerberg/building-global-community/10154544292806634.

191. Sofia Grafanaki, *Platforms, the First Amendment and Online Speech: Regulating the Filters*, 39 PACE L. REV. 111 (2018).

the promotion and suppression of certain pieces of content ("prioritization" by the platform), the impact of which is not limited to individual pieces of content but to the entire flow of content on the platform. This measure enables a platform to increase the popularity and impact of highly visible content, while marginalizing and limiting the impact of other content. The social media companies decide what is available to whom and on what basis. "Changes in the prominence regime could be used to manipulate media,"[192] and this signifies a threat to freedom of expression. Down-ranking (referred to by a telling euphemism as "curation") of speech that is otherwise relevant for public deliberation is an extremely dangerous tool in the hands of social media platforms.[193] All of this is done with the aim of providing personalized services and serving individual user needs (as guessed by the platform), relying on information collected about each and every user, their previous online presence, and their platform-generated profile. Thus, each user unknowingly, and indeed without explicit consent, influences the content of the service he or she receives, while the platform actively exerts an influence over the user's intentions and is capable of influencing the user. The resulting consequences have an impact on the decisions users make as consumers, and also on the discussion of public affairs, access to information, and the diversity of opinion—in other words, the quality of the democratic public sphere.

The real power of a platform to influence the discussion of public affairs is not rooted in the capacity to remove individual pieces of content or ban users. Platforms use algorithms that enable them, on the basis of data collected about each user, to personalize each and every piece of content accessible to and consumed by their users. An obvious example is Facebook's newsfeed, which includes only a small portion of all content published by a user's acquaintances and the pages he or she follows. Obviously, this practice is also justified by practical considerations, since the platform serves its users by keeping the content available to them organized in some way and by showing them the content in which they are most likely to be interested. However, we do not exactly know the basis on which the platform relies when sorting such content, how it tries to find an appropriate balance between public

192. Eleonora M. Mazzoli & Damian Tambini, *Prioritisation Uncovered: The Discoverability of Public Interest Content Online*, COUNCIL OF EUROPE STUDY DGI (Nov. 2020) 19, 42 https://rm.coe.int/publication-content-prioritisation-report/1680a07a57.

193. Sarah C. Haan, *Facebook and Politicians' Speech*, 70 AM. U. LAW REV. F. 203 (2021). The presence of similar recommendation systems on other online platforms should also be noted, see the example of music streaming platforms: Tamás Tófalvy & Júlia Koltai, *"Splendid Isolation": The Reproduction of Music Industry Inequalities in Spotify's Recommendation System*, NEW MEDIA & SOC'Y. 1 (2021).

issues and holiday photos, and what the business interests are or could be behind featuring certain specific pieces of content. It should be noted that not all social media platforms edit user content as comprehensively as Facebook, but each platform attempts to show pieces from all potentially available content to a user that is most likely to meet his or her interests.

An issue that arises in relation to the compilation of a news feed and the pieces of content shown to a user in general is whether it can be considered as protected speech by the platform. While such an argument would seem difficult to maintain in the context of filtering and removing content, it could be possible that the compilation of a news feed does eventually produce some kind of content that is new and did not exist before, the individual components of which were not produced or commissioned by the platform, but where the work of compilation was indeed performed by the platform according to its own decisions and considerations. On the one hand, if such a compilation is protected under the freedom of speech, it would be difficult to influence it from the outside. On the other hand, if the compilation is considered similar to the editing activities of the traditional media, it might be possible to apply the rules and doctrines of such media with some reasonable adjustments. In the words of Robin Foster:

> There are no exact parallels for the new digital intermediaries identified here—most are not neutral "pipes" like ISPs, through which all Internet content flows (although Twitter is close to this); nor are they pure media companies like broadcasters or newspapers, heavily involved in creative and editorial decisions. But they do perform important roles in selecting and channelling information, which implies a legitimate public interest in what they do.[194]

Paul Bernal even calls the supposed neutrality of gatekeepers (Facebook among them) a "myth."[195] The selection of content is not a neutral or value-neutral activity, and it reflects numerous interests of a platform. This might not be a problem in and of itself, but it raises concerns that users are not familiar with those interests and values. Users cannot really know (or it would take extreme effort on their side to find out) what else is really out there apart from the content they are shown. As Klonick has highlighted, it is not a priority for social media platforms at this time to ensure adequate opportunities for all to participate in public discourse.[196]

194. ROBIN FOSTER, NEWS PLURALITY IN A DIGITAL WORLD 30 (2012).

195. Bernal, *supra* note 188, at 71–101.

196. Klonick, *supra* note 71, at 1665.

It seems that the architecture of platforms, the editorial decisions governing their general operations and their underlying values are more important than any individual decision made by a platform in response to a notice concerning a given piece of content, as those factors determine the overall functioning of the platform and have an impact on the ability of all the users to access information.[197] The provision of a personalized service to users could act to suppress their overall picture of news and information on public affairs. This means that the number of news reports produced by traditional media can be drastically reduced at the whim of Zuckerberg[198] which is also announced by the platform from time to time.[199] When a bill unfavourable to Facebook was published in early 2021, it banned Australian news providers and media companies overnight from posting new content and other users from sharing content from such companies.[200] The draft required platforms to pay content producers if their content is made available by the platform. After a few weeks of discussion, an agreement was reached, after which the content of these providers became accessible again.[201] According to documents leaked by a whistleblower in September 2021, Facebook shielded millions of VIP users from standard moderation protocols, by using a program that whitelisted millions of VIP users from the company's standard content moderation practices, despite the company's insistence that all rules apply equally to all users.[202] According to the leaked documents, Facebook changed content policies for several weeks surrounding the 2020 U.S. elections, when the

197. Frank Fagan, *Systemic Social Media Regulation*, 16 DUKE L. & TECH. REV. 393 (2018).

198. Emily Bell, *Why Facebook's News Feed Changes are Bad News for Democracy*, THE GUARDIAN (Jan. 21, 2018), https://www.theguardian.com/media/media-blog/2018/jan/21/why-facebook-news-feed-changes-bad-news-democracy.

199. Kevin Roose & Mike Isaac, *Facebook Dials Down the Politics for Users*, THE NEW YORK TIMES (Feb. 10, 2021), https://www.nytimes.com/2021/02/10/technology/facebook-reduces-politics-feeds.html.

200. Archie Bland, *Facebook Over-Enforced Australia News Ban, Admits Nick Clegg*, THE GUARDIAN (Feb. 24, 2012), https://www.theguardian.com/technology/2021/feb/24/facebook-over-enforced-australia-news-ban-admits-nick-clegg.

201. Kari Paul, *What Facebook's Australia News Ban could Mean for its Future in the US*, THE GUARDIAN (Feb. 27, 2021), https://www.theguardian.com/technology/2021/feb/27/facebook-australia-news-ban-us-legislation.

202. Jeff Horwitz, *Facebook Says Its Rules Apply to All. Company Documents Reveal a Secret Elite That's Exempt*, THE WALL STREET JOURNAL (Sept. 13, 2021), https://www.wsj.com/articles/facebook-files-xcheck-zuckerberg-elite-rules-11631541353?mod=djemalertNEWS.

company gave lower priority to political content on its news feed, but after the election the platform soon went back to algorithms that valued engagement over all else, realizing that "if they change the algorithm to be safer, people will spend less time on the site, they'll click on less ads, and [Facebook] will make less money."[203]

The tools available to platforms for influencing the public are downright frightening, and given that, for many users, social media platforms have become the primary or often the only source of news, restricting public interest content in no way improves the quality of democratic decision-making.

Facebook's reasons for changing how algorithms present (or hide) certain types of content are predominantly financial, not a matter of principle. Facebook is in fact in competition with such media, even if it does not produce any content, and the company seeks to maximize its revenues from those media in return for presenting their content to its users. The provision of personalized services and news services in particular can reduce the diversity and selection of news that individual users come across when using the platform. In the era of traditional media, it was inevitable for readers to see content they did not specifically look for or agree with, but the comfort of personalization eliminates this unpleasantness. The personalization of news goes against the very concept of the marketplace of ideas, as users do not meet opinions that contradict their own personal views and opinions unless they specifically look for them.[204]

A platform may also interfere with its news feed in line with its political views and social objectives, and this very capacity poses a direct threat to the public sphere and the democratic expression of opinions. The existence of this phenomenon was demonstrated by a scandal in 2016 when tech-blog Gizmodo reported allegations from Facebook staff members that the company suppressed conservative topics and sources deliberately and in a systemic manner. The platform had claimed previously that the content of Trending Topics (a service listing topics that are most actively discussed by other users of the platform, also known as "hot topics") was compiled by algorithms exclusively on the basis of actual user activity and without any direct human intervention. However, former employees of the company

---

203. Kari Paul & Dan Milmo, *Facebook Putting Profit Before Public Good, Says Whistleblower Frances Haugen*, THE GUARDIAN (Oct. 3, 2021), https://www.theguardian.com/technology/2021/oct/03/former-facebook-employee-frances-haugen-identifies-herself-as-whistleblower.

204. Sarah Eskens, Natali Helberger & Judith Moeller, *Challenged by News Personalisation: Five Perspectives on the Right to Receive Information*, 17 J. MEDIA L. 259, 281 (2017).

reported that they had to select the topics on the basis of political considerations.

According to reports, links to certain conservative websites were not allowed in the Trending Topics section, even if they were among the most frequently shared content on the platform.[205] The case showed clearly that the selection of pieces of news that were to be featured and widely discussed on Facebook was not influenced by neutral algorithms but human editors (known as "news curators").[206] In essence, the scandal resulted in the defeat of an important taboo and a paradigm shift regarding the role of the platform. Facebook became an actual news editor and, as such, similar to traditional media.[207] Even though Trending Topics was phased out by the platform eventually, it seems hard not to believe that similar news editing practices might be used by other services of the platform or outside the U.S. [208] If Facebook is considered a news editor, it might just be reasonable to extend the scope of legal provisions applicable to the news editors of the legacy media to social media platforms.

The Cambridge Analytica scandal became a global issue. The data analytics firm worked with Trump's election team and for the Brexit campaign and harvested the personal data of up to eighty-seven million Facebook users (U.S. voters), in the platform's biggest known data breach so far. Cambridge Analytica used these data to build a software program to predict and influence choices at the ballot box. The profiling

---

205. Philip Bump, *Did Facebook Bury Conservative News? Ex-Staffers Say Yes*, THE WASHINGTON POST (May 9, 2016), https://www.washingtonpost.com/news/the-fix/wp/2016/05/09/former-facebook-staff-say-conservative-news-was-buried-raising-questions-about-its-political-influence/?noredirect=on&utm_term=.8338634b2401; Michael Nunez, *Former Facebook Workers: We Routinely Suppressed Conservative News*, GIZMODO (May 9, 2016), https://gizmodo.com/former-facebook-workers-we-routinely-suppressed-conser-1775461006.

206. Sam Thielman, *Facebook News Selection is in Hands of Editors Not Algorithms, Documents Show*, THE GUARDIAN (May 12, 2016), https://www.theguardian.com/technology/2016/may/12/facebook-trending-news-leaked-documents-editor-guidelines.

207. Natali Helberger & Damian Trilling, *Facebook is a News Editor: The Real Issues to be Concerned About*, LSE MEDIA POLICY PROJECT (May 26, 2016), http://blogs.lse.ac.uk/mediapolicyproject/2016/05/26/facebook-is-a-news-editor-the-real-issues-to-be-concerned-about.

208. Chris Morris, *Facebook Kills "Trending" Topics, Will Test "Breaking News" Label*, FORTUNE (Jun. 1, 2018), http://fortune.com/2018/06/01/facebook-kills-trending-news-topics.

of U.S. individuals was aimed at targeting them with personalized political advertisements, thereby influencing several elections.[209]

In the last weeks of the 2020 presidential campaign, the *New York Post* published a story that alleged that while Joe Biden (the then-future President) was Vice President of the U.S., he had engaged in corrupt activities relating to the employment of his son Hunter Biden by the Ukrainian gas company Burisma. The story's credibility is still disputed. However, according to established facts, Hunter Biden was indeed hired by Burisma, and he received large sums of money from the company, but corruption on the part of the then Vice President remains unproven.[210] After the publication of the story on the *New York Post*'s website, Twitter and Facebook implemented measures to prevent the sharing of the article. Later on, they explained their decision as forming part of their ongoing struggle against fake news.[211] In this instance, the "suppression is a bigger scandal than the actual story."[212] Commentators accused the platforms of censorship, partisanship, double standards, and the intention to influence the presidential elections.[213]

The activities of gatekeepers raise questions concerning both their possible direct interference with the freedom of speech and also the issue of media (or, in the case of social media platforms, content) pluralism or diversity, which is one of the main objectives of media regulation in Europe. The regulatory regimes aim to increase the diversity of published content and opinions concerning public matters regarding the television and radio markets. To this end, provisions pertaining to content and structural restrictions seeking to prevent the concentration of ownership may also be adopted to a limited extent, and the public service media can also work to expand the media offering. Given the market influence of gatekeepers, it seems reasonable to raise

209. Alvin Chang, *The Facebook and Cambridge Analytica Scandal, Explained with a Simple Diagram*, VOX (Mar. 23, 2018), https://www.vox.com/policy-and-politics/2018/3/23/17151916/facebook-cambridge-analytica-trump-diagram.

210. Katie Robertson, *New York Post Published Hunter Biden Report Amid Newsroom Doubts*, THE NEW YORK TIMES (Oct. 18, 2020), https://www.nytimes.com/2020/10/18/business/media/new-york-post-hunter-biden.html.

211. Kelvin Chan, *Twitter CEO Says it was Wrong to Block Links to Biden Story*, AP NEWS (Oct. 16, 2020), https://apnews.com/article/business-media-social-media-censorship-ec529ef85c1e72cefe0ae9450e118b9c.

212. Matt Taibbi, *With the Hunter Biden Expose, Suppression Is a Bigger Scandal Than the Actual Story*, TK NEWS (Oct. 24, 2020), https://taibbi.substack.com/p/with-the-hunter-biden-expose-suppression-136?fbclid=IwAR0Sbma_0gfwmaK95xsdb_11hDYF-2Kb_jLV4VXeu5S3NJr2_PuQcRmbge8.

213. Glenn Greenwald, *My Resignation from the Intercept*, GREENWALD (Oct. 29, 2020), https://greenwald.substack.com/p/my-resignation-from-the-intercept.

the issue of media pluralism, a matter that is quite common with regard to television and radio services, and the possible adaptation of known media regulatory solutions to the online environment. [214]

Note the strange paradox here: in the beginning, the Internet was heralded as the ultimate solution to the scarcity of content and promised to render earlier forms of media regulations meaningless, while we are now facing problems already all too well-known from the media market, but on a much larger scale.

Back in 1996, Owen Fiss argued that freedom of speech had traditionally been considered as a shield from interference by the state, thereby protecting individual freedom. This doctrine fostered a media ecosystem that benefited large media corporations, and gave them powers over the public sphere, silencing individuals who did not possess any opinion-shaping power or any opportunity to express themselves in the public sphere. This was the irony of free speech.[215] Moran Yemini calls the new digital ecosystem of public communication the "new irony" of free speech.[216]

Nevertheless, freedom of speech remains primarily a "negative right" that protects the content of opinions against external sources of interference, but it does not guarantee the right to use platforms that facilitate the effective exercise of this freedom. Similar to traditional press and media regulations, a recognized right to reach an audience does not exist in the context of the Internet either.[217]

2. The Media Regulation Toolbox

If social media platforms were primarily to be seen by regulators as media services, the application of legacy media obligations would naturally arise. However, the media regulation approaches common in Europe cannot be applied without addressing the content diversity issue posed by social media platforms. Even so, adopting some elements of certain regulatory instruments, after suitable modification, adapting them to the operational characteristics of the platforms, would not be unthinkable. The following section reviews this set of media regulation tools, including the difficulties that may arise in their application. Rules

214. Robin Mansell, *The Public's Interest in Intermediaries*, 17 J. POL'Y. REGUL. STRATEGY TELECOMM. INFO. MEDIA 8 (2015); Natali Helberger et. al., *Regulating the New Information Intermediaries as Gatekeepers of Information Diversity*, 17 INFO 50 (2015).

215. OWEN M. FISS, THE IRONY OF FREE SPEECH (1996).

216. Moran Yemini, *The New Irony of Free Speech*, 20 COLUM. SCI. & TECH. L. REV. 119 (2018).

217. Jennifer A Chandler, *A Right to Reach an Audience: An Approach to Intermediary Bias on the Internet*, 35 HOFSTRA L. REV. 1095, 1096–1103 (2007).

to increase supply diversity are organized around the idea of media pluralism.

The regulation of television, radio, and on-demand audiovisual and radio media services seek to remedy, through indirect means, distortions in public communication caused by partisanship, the lack of diversity, or the publication of false information. Generally speaking, these tools try to accommodate as many different opinions as possible in the debate on public affairs. In line with the theoretical requirements of media pluralism, the entire media market should collectively cater for the diversity of opinions and available content and establish a balance between them.[218] This requirement primarily imposes tasks on the state in respect of the regulation of the media market, and in practice such regulation mainly concerns traditional television and radio broadcasting. However, since the market of social media platforms tend to be monopolistic, overseeing the market by state agencies does not seem to be helpful for maintaining and developing the public sphere.

Individual legal systems try to achieve the objective of media pluralism primarily by controlling media concentration. Such rules seek to prevent, by restricting ownership, the emergence of media market concentration.[219] In recent years, the European Commission has repeatedly attempted to take action against, for example, market abuses by companies such as Google.[220] In addition to the powerful voices that even suggested breaking up these giants, there were also worldwide doubts as to whether the restriction of concentration could provide an appropriate response to the problems posed by the platforms, and whether strong intervention in market conditions would do more harm to publicity than it would prevent.[221]

Based on the right of reply, access to the content of a media service provider is granted in response to content published previously by the service provider. Article 28 of the AVMS Directive[222] prescribes that E.U. Member States should introduce national legal regulations with regard to television broadcasting that ensure adequate legal remedies for those whose personality rights have been infringed through false statements. Such regulations are known Europe-wide and typically

218. *See* EWA KOMOREK, MEDIA PLURALISM AND EUROPEAN LAW (2012).

219. EDWIN C. BAKER, MEDIA CONCENTRATION AND DEMOCRACY: WHY OWNERSHIP MATTERS (2006).

220. Jeanne Whalen, *Europe Fined Google Nearly $10 Billion for Antitrust Violations, But Little Has Changed*, THE WASHINGTON POST (Nov. 10, 2020), https://www.washingtonpost.com/technology/2020/11/10/eu-antitrust-probe-google.

221. Steven C. Salop, *Dominant Digital Platforms: Is Antitrust Up to the Task?*, 130 YALE L. J. F. 563 (2021).

222. AVMS Directive, *supra* note 13.

impose obligations on the printed and online press alike.[223] It is important to highlight that the function of the right of reply is twofold. On the one hand, it serves the protection of the personality rights (the reputation or honor) of the person attacked. On the other hand, it serves the right of the public to appropriate, truthful information.

The compatibility of the right of reply and Article 10 of the Convention has been confirmed in several decisions of the European Court of Human Rights.[224] In *Melnychuk v. Ukraine*,[225] the court established that the right of reply constituted a part of the freedom of speech of the applicant. Thus, rather than limiting the freedom of the press of the publisher of the newspaper carrying the injurious content, the opposite is true. The right is an instrument that enables the complainant to effectively exercise their freedom of speech in the forum where the complainant has been attacked. In *Kaperzynski v. Poland,*[226] the European Court of Human Rights held:

> The Court is of the view that a legal obligation to publish a rectification or a reply may be seen as a normal element of the legal framework governing the exercise of the freedom of expression by the print media . . . . Indeed, the Court has already held that the right of reply, as an important element of freedom of expression, falls within the scope of Article 10 of the Convention. This flows from the need not only to be able to contest untruthful information, but also to ensure a plurality of opinions, especially on matters of general interest such as literary and political debate.[227]

The first major decision by the Supreme Court of the United States concerning a right of reply law was *Red Lion Broadcasting v. Federal Communications Commission*.[228] It examined the constitutionality of the Federal Communications Commission's fairness doctrine, which required that some discussion of public issues must be presented on broadcast stations, and that each side of those issues must be given fair coverage. It contained a specific right of reply element: if, during the presentation of a controversial issue, an attack was made, "upon the honesty, character, integrity or like personal qualities of an identified

---

223. Kyu Ho Youm, *The Right of Reply and Freedom of the Press: An International and Comparative Perspective*, 76 GEO. WASH. L. REV. 1017 (2008); András Koltay, *The Right of Reply in a European Comparative Perspective*, HUNGARIAN J. OF LEGAL STUD.—ACTA JURIDICA HUNGARICA 73 (2013).

224. Ediciones Tiempo S.A. v. Spain, no. 13010/87, decision of 12 July 1989.

225. Melnychuk v. Ukraine, no. 28743/03, decision of 5 July 2005.

226. Eur. Ct. H.R. Kaperzynski v. Poland, no. 43206/07, judgment of 3 April 2012.

227. *Id*. at para 66.

228. 395 U.S. 367 (1969).

person or group," the attacked person must be given an opportunity to reply. The same obligation applied if a political candidate's views were endorsed or opposed, which entailed the broadcaster giving the opposing candidate or the opponents of the endorsed candidate the opportunity to respond. The Court unanimously upheld the regulations.

It came as a slight surprise in the light of *Red Lion* that, only five years later, the Court—again unanimously—struck down a piece of Florida legislation that required the printed press to give the right of reply to candidates for political office who had been assailed over their personal character or official record.[229] The Court ruled in favour of the autonomous press:

> [T]he implementation of a remedy such as an enforceable right of access necessarily . . . brings about a confrontation with the express provisions of the First Amendment . . . . Compelling editors or publishers to publish that which "'reason' tells them should not be published" is what is at issue in this case. The Florida statute operates as a command in the same sense as a statute or regulation forbidding [the newspaper] to publish specified matter . . . . [T]he Florida statute fails to clear the barriers of the First Amendment because of its intrusion into the function of editors.[230]

Many commentators celebrated the *Miami Herald Publishing Co. v. Tornillo* decision as a victory for press freedom and blamed the Court for the serious mistake it made in *Red Lion Broadcasting Co. v. Federal Communications Commission*.[231] They argued that a free media market, even with its considerable failings, is always better than one that is regulated by the state. Other authors celebrate *Red Lion and* hold that *Miami Herald* was wrong. For them, ensuring that people are presented with a wide range of views about public issues is necessary to make democracy work, and this aim can justify state intervention.[232]

---

229. Miami Herald Publishing Co. v. Tornillo, 418 U.S. 241 (1974).

230. *Id*. at 254–56, 258.

231. For just a handful of examples from the vast amount of literature, *see* Kenneth A. Karst, *Equality as a Central Principle in the First Amendment*, 43 U.C.L.R. 20 (1975–76); Vincent Blasi, *The Checking Value in First Amendment Theory*, 2 AMERICAN BAR FOUNDATION RESEARCH JOURNAL 521 (1977); EDWIN C. BAKER, HUMAN LIBERTY AND FREEDOM OF SPEECH (1989).

232. Cass Sunstein, *A New Deal for Speech*, 17 HASTINGS COMM. AND ENT. L. J. 137 (1994); Charles W. Logan, *Getting Beyond Scarcity: A New Paradigm for Assessing the Constitutionality of Broadcast Regulation*, 85 CAL. L. REV. 1687 (1997); Jerome A. Barron, *Rights of Access and Reply to the Media in the United States Today*, 25 COMM. AND THE LAW 1 (2003).

In certain European states, the regulation promoting media pluralism also includes the requirement for impartial news coverage, on the basis of which public affairs need to be reported impartially in programs providing information on them. Such regulation may apply to television and radio broadcasters, and it has been implemented in several states in Europe.[233] Those who argue against maintaining this rule point out that, since the former scarcity of information has been eliminated, and hence, in this new media world, everyone can obtain information from countless sources, the earlier regulatory models have become redundant or, one might say, anachronistic. By contrast, as Steven Barnett noted, for as long as television journalism can be differentiated from Internet journalism, there is no reason to stop having media-specific rules.[234] Mike Feintuck argued that the earlier assumption, suggesting that in a free and unrestricted media market a diversity of opinions would automatically appear and hence impartiality would arise, has proven unfounded.[235] As Richard Sambrook put it, "[i]f the words 'impartiality' and 'objectivity' have lost their meanings, we need to reinvent them or find alternative norms to ground journalism and help it serve its public purpose—providing people with the information they need to be free and self-governing."[236]

Safeguarding media pluralism can impose obligations not only on the state and media service providers, the publishers of printed and online press products, but also on those distributing television and radio programs (cable and satellite broadcasters). Pursuant to the must carry rules, distributors need to include the programs of certain broadcasters in the services broadcast to audiences, which means that they need to allocate a certain part of the distribution capacity to certain broadcasters, typically public service or local broadcasters, in order to safeguard media pluralism in the interest of the public.

The must carry restriction was also recognized by the Supreme Court of the United States as constitutional in the *Turner Broadcasting*

---

233. *See*, *e.g.*, the German regulations (Rundfunkstaatsvertrag, §§ 25–34); U.K. regulation §§ 319(2)(c); 319(2)(d), 319(8); 320 of the Communications Act 2003; and § 5 of the Broadcasting Code; *see also* the recent decision of the ECtHR, Associazione Politica Nazionale Lista Marco Pannella v. Italy, no. 66984/14, judgment of 31 August 2021.

234. Steven Barnett, *Imposition or Empowerment? Freedom of Speech, Broadcasting and Impartiality*, FREEDOM OF EXPRESSION AND THE MEDIA 58 (Merris Amos, Jackie Harrison and Lorna Woods eds., 2012).

235. Mike Feintuck, *Impartiality in News Coverage: The Present and the Future*, FREEDOM OF EXPRESSION AND THE MEDIA 88 (Merris Amos, Jackie Harrison and Lorna Woods eds., 2012).

236. RICHARD SAMBROOK, DELIVERING TRUST: IMPARTIALITY AND OBJECTIVITY IN THE DIGITAL AGE 39 (2012).

System, Inc. v. Federal Communications Commission cases.[237] Justice Kennedy, drafting the reasoning in the first case, established that the constitutional restriction of cable services is easier than that of broadcasting, because the former is not expressly related to the restriction on content. However, more in-depth analysis is required to establish whether or not the legal restriction is really necessary and justified. The second *Turner* decision answered this question in the affirmative. Sunstein considered this decision of the Court to be a sort of overture to the new, community-based interpretation of the First Amendment, as a small majority of the board recognized that a strictly content-neutral restriction of access to media is constitutional, as it is in the public interest.[238]

In certain cases, media regulation may require media service providers, as a condition of their entitlement to provide media services, to publish information of public interest, to provide local news, and to reserve a certain proportion of airtime for public service programs.[239] In addition, public service media providers are strong players in the media market in European states, primarily operating by using public financial resources.[240]

3. Applying the Principles of Media Regulation to Social Media Platforms

The media regulation solutions described above would certainly not work without any changes taking place in the world of social media. These rules were adopted in the era of technological scarcity, and it would not be possible to justify their strict application to today's public platforms, characterized by excessive abundance of content.[241] However, the considerations underlying these rules have not become void. The European view maintains that it is not the principles of media regulation that are in decline, but at most its methods in a public sphere dominated by Internet communication and platforms. The algorithms of the platforms may also draw attention to the existence of different opinions and the platform can provide an opportunity for a person whose reputation has been violated to respond to false factual

237. *Turner I*, 512 U.S. 622 (1994); *Turner II*, 520 U.S. 180 (1997).

238. Cass R. Sunstein, *The First Amendment in Cyberspace*, 104 YALE L. J. 1757, 1765–81 (1995).

239. *See e.g.*, the U.K.'s Communications Act 2003, s. 287(2).

240. KAREN DONDERS, PUBLIC SERVICE MEDIA IN EUROPE: LAW, THEORY AND PRACTICE (2021).

241. *See e.g.*, Mike Jayne, *Fairness Doctrine 2.0: The Ever-Expanding Definition of Neutrality Under the First Amendment*, 16 FIRST AMEND. L. REV. (2018) (with regard to the fairness doctrine's possible application to online platforms).

statements. In principle, the platform may be required to present a certain amount or proportion of news and information of public interest to its users.

The kind of diversity that can be expected at all from social media also differs from the requirements applicable to legacy media, in that it does not concern the production or commissioning of diverse content but would instead envisage the appropriate selection of already diverse user content and make a sufficiently diverse flow of information accessible to individual users. Natali Helberger, Kari Karppinen, and Lucia D'Acunto call this exposure diversity,[242] and consider it conceivable that, if this is integrated in the operation of the platforms from the outset (diversity by design), platforms can deliver on the need for diversity.[243]

Even though the concept of "due impartiality in news coverage," a requirement under traditional media regulation, could serve as an appropriate starting point for introducing a new regulatory scheme, a legal system may not require social media platforms to operate with the same degree of impartiality as a television or radio news program, particularly because (i) a platform does not produce any content that would be relevant in this context, and (ii) not even a platform is capable of overseeing the entire body of content generated by its users. Requiring a platform to attempt to present content in an entirely impartial manner would mean that it is subject to the same obligations as a television or radio editor, despite the above-mentioned characteristics. The regulation of electronic program guides seems to be a more appropriate analogy, as it requires service providers to present certain important pieces of (public service media) content in a distinctive manner. This obligation may be labelled by various names, such as "findability," "due prominence," or even also as "exposure diversity."[244]

Encountering a wide variety of content, including some that the user would not have deliberately sought, such as reactions to a false statement of fact, or raising awareness of dissenting opinions, is good for democratic publicity and contributes to informed opinion-making, thus enhancing the quality of democratic decision-making and counteracting the oft-cited "echo chamber" or "filter bubble" effect. According to some theories, the Internet has a negative impact on

242. Natali Helberger et. al., *Exposure Diversity as a Design Principle for Recommender Systems*, 21 INFO. COMMC'N. SOC'Y. 191 (2018).

243. *Id*. at 203–04.

244. Bart van der Sloot, *Walking a Thin Line: The Regulation of EPGs*, 3 JIPITEC 138 (2012).

various social groups and their members. Cass Sunstein warned us of the dangers of fragmentation as early as two decades ago.[245] First, every user can decide individually which content to read, view, or follow. This might prompt users to prefer forums that reinforce and resonate with their existing opinions, and which typically provide positive feedback, without having to face others holding an opposing opinion. Second, social media amplifies this phenomenon by delivering a personalized news stream to each user, which consists of content originating from friends with similar opinions and media outlets preferred by the user. This is the *Daily Me*, a form of news source that is always in agreement with the reader, thereby intensifying his or her pre-existing liberal or conservative views and opinions.[246] These customized services create the "filter bubble effect," that is, they trap users in a circle of content that is identical to or consistent with their own views and mostly hide other content from them.[247] On the other hand, traditional media compile content on their own without any input from the reader or viewer, making it inevitable that members of the audience will be confronted with various points of view. The benefits of this approach include the emergence of a more complex worldview and the reinforcement of critical thinking.[248] In contrast, the dominance of social media and search engines deepens the gap between individuals with conflicting opinions, thereby weakening social cohesion and strengthening extremism (polarization).[249] Other researchers seek to disprove Sunstein's theory and argue that "omnivore" Internet users are the rule, while calling for the previous world of media to be presented in a more critical manner.[250]

Transparency of prioritization by platforms, namely ensuring that the users concerned and the public bodies supervising the operation are aware of what is happening on the platform, can be required in principle (this does not necessarily mean full transparency of the operation of algorithms in a technological sense, which is difficult or even impossible to implement). Some European countries have already

245. CASS R. SUNSTEIN, REPUBLIC.COM (2001); CASS R. SUNSTEIN, REPUBLIC.COM 2.0 (2007).

246. A term coined by Nicholas Negroponte and also used by Sunstein, *see* Cass R. Sunstein, ch 1, *in* #REPUBLIC. DIVIDED DEMOCRACY IN THE AGE OF SOCIAL MEDIA (2017).

247. ELI PARISER, THE FILTER BUBBLE: WHAT THE INTERNET IS HIDING FROM YOU (2012) (in the context of search engines).

248. Sunstein, *supra* note 239, at 140–48.

249. *Id.* at ch. 3.

250. TIM WU, MASTER SWITCH 214–15 (2010); Matthew Gentzkow & Jesse M. Shapiro, *Ideological Segregation Online and Offline* (Chicago Booth Research Paper No. 10-19, 2010).

started reflecting on how to require platforms to offer diverse content and to prevent bias in publicity.[251] Of course, state intervention is not without risk and may lead to state censorship in the name of acting against censorship of platforms, which regulatory suggestions must keep in mind.

The regulation of "recommender systems" is also on the E.U.'s agenda. The Digital Services Act proposes making their operation more transparent and increasing the importance of user preferences:

> Article 29, Recommender Systems
>
> 1. Very large online platforms that use recommender systems shall set out in their terms and conditions, in a clear, accessible and easily comprehensible manner, the main parameters used in their recommender systems, as well as any options for the recipients of the service to modify or influence those main parameters that they may have made available, including at least one option which is not based on profiling, within the meaning of Article 4 (4) of Regulation (EU) 2016/679.
>
> 2. Where several options are available pursuant to paragraph 1, very large online platforms shall provide an easily accessible functionality on their online interface allowing the recipient of the service to select and to modify at any time their preferred option for each of the recommender systems that determines the relative order of information presented to them.

The latest legislation on this area in Germany aims to help achieve one of the main objectives of media regulation affecting platforms.[252] The regulation obliges social media platforms, video sharing platforms, and search engines to be non-discriminatory in terms of content and to prioritize public service content, while not restricting user preferences.[253] These provisions aim to promote the diversity of content on platforms.

---

251. Paddy Leerssen, *The Soap Box as a Black Box: Regulating Transparency in Social Media Recommender Systems*, 11 EUROPEAN J. L. TECH. (2020).

252. Natali Helberger, Paddy Leerssen & Max Van Drunen, *Germany Proposes Europe's First Diversity Rules for Social Media Platforms*, MEDIA POLICY PROJECT (May 29, 2019), https://blogs.lse.ac.uk/mediapolicyproject/2019/05/29/germany-proposes-europes-first-diversity-rules-for-social-media-platforms.

253. Jan C. Kalbhenn, *New Diversity Rules for Social Media in Germany*, CENTER FOR MEDIA PLURALISM AND MEDIA FREEDOM, (Feb. 21, 2020), https://cmpf.eui.eu/new-diversity-rules-for-social-media-in-germany.

V. CONCLUSIONS

The impact of social media platforms on freedom of speech is extremely wide-ranging. Social and legacy media are similar in terms of editorial activity, but there are also significant differences between them. Accordingly, media regulation cannot be applied to platforms without any change. The approach used in Europe, according to which, in certain cases, the lawfulness of content must be decided by platforms, raises concerns in terms of freedom of speech. At the same time, it is also clear that the judicial system or a public authority would not be able to handle the workload associated with the operation of the platforms, so the notice-and-takedown system remains the basis for the liability of the platforms as a kind of emergency measure.

The contract between a platform and its users provides an opportunity to take the interests of users regarding freedom of speech more into account than is currently the case. State regulation can also help in this regard. The platforms themselves are private actors, which claim a right to protect their own freedom of speech, so this must be taken into account in any regulatory attempts.

The regulations and proposals that have been made so far in European states primarily encourage stronger action against harmful, dangerous content and, accordingly, are less concerned with the protection of freedom of speech. The proposed regulation of the European Union (the Digital Services Act) and some ideas at national level also aim to strengthen users' freedom of speech, mainly through the introduction of appropriate procedural guarantees restricting the scope of private regulation and the creation of an independent forum for redress.

Another aspect of editing by platforms, the regulation of prioritization between content, is currently on the agenda but has been accorded less emphasis. The diversity of content that is actually available to users and easy access to public interest content is a fundamental concern of democratic publicity. The media regulation solutions already widely used in Europe may inspire the regulation of platforms, and the principles and values underlying regulation will not melt into thin air as a result of technological progress.