þÿFrom Delfi to Sanchez when can an online com platform be responsible for third-party comments? : An analysis of the practice of the ECtHR and some reflections on the Digital Services Act

Korpisaari (ex. Tiilikka), Päivi

2022

þÿKorpisaari (ex. Tiilikka), P 2022, 'From Delfi to Sanchez when can communication platform be responsible for third-party comments? An analysis of the practice of the ECtHR and some reflections on the Digital Services Act', The Journal of Media Law, vol. 14, no. 2, pp. 352-377. https://doi.org/10.1080/17577632.2022.2148335

http://hdl.handle.net/10138/356247
https://doi.org/10.1080/17577632.2022.2148335

Routledge
Taylor & Francis Group

Journal of Media Law

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rjml20

From *Delfi* to *Sanchez* – when can an online communication platform be responsible for third-party comments? An analysis of the practice of the ECtHR and some reflections on the Digital Services Act

Päivi Korpisaari

To cite this article: Päivi Korpisaari (2022) From *Delfi* to *Sanchez* – when can an online communication platform be responsible for third-party comments? An analysis of the practice of the ECtHR and some reflections on the Digital Services Act, Journal of Media Law, 14:2, 352-377, DOI: 10.1080/17577632.2022.2148335

To link to this article: https://doi.org/10.1080/17577632.2022.2148335

Published online: 24 Nov 2022.

JOURNAL OF MEDIA LAW
2022, VOL. 14, NO. 2, 352–377
https://doi.org/10.1080/17577632.2022.2148335

Routledge
Taylor & Francis Group

From *Delfi* to *Sanchez* – when can an online communication platform be responsible for third-party comments? An analysis of the practice of the ECtHR and some reflections on the Digital Services Act

Päivi Korpisaari

Communication Law at the Faculty of Law, University of Helsinki, Helsinki, Finland

ABSTRACT

While social media services offer a useful platform for obtaining information as well as presenting and commenting on opinions, people can still be silenced by fear of hate speech and insults on the Internet. As a result, the expanded freedom of expression can also reduce the range of opinions and information. This article identifies and analyses the conditions under which online communication platform administrators can be held liable for user-generated content. The focus is on the criteria laid down by the ECtHR in recent cases. The outcome is that liability is exceptional, arising mainly in cases of inciting hatred and violence. Although the Digital Services Act, with its notice-and-action mechanism, offers a cheaper, faster, and often more effective way of reducing insulting and defamatory speech than court proceedings, the impact of the mechanism on freedom of expression and freedom to conduct business must be considered.

ARTICLE HISTORY Received 4 January 2022; Accepted 24 October 2022

KEYWORDS Freedom of expression; hate speech; liability; platform; digital services act

Scope and aim of research

Towards increased regulation of the internet

Today, Europe-wide reflection is focused on how to create effective measures to reduce and minimise the effects of hate speech. For example, Germany,[1]

CONTACT Päivi Korpisaari paivi.korpisaari@helsinki.fi

[1]For a closer examination, see Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken (*Facebook-Gesetz*), 2017. See also M Hong, 'Regulating hate speech and disinformation online while protecting freedom of speech as an equal and positive right – comparing Germany, Europe and the United States' (2022) 14 Journal of Media Law 76.

France,[2] and Austria[3] have adopted legislation requiring certain online platforms to remove illegal content. An Online Safety Bill has been introduced in Britain, but it is not certain whether the act will pass in Parliament.[4]

Illegal speech on the Internet has also been a topical issue at the level of the European Union. On 1 March 2018 the European Commission issued a recommendation on measures to tackle illegal content online.[5] The Recommendation is aimed at hosting service providers rather than actual online communication platforms. The Commission considers that rapid removal of illegal content is an effective measure to limit wider dissemination and harm caused to those subjected to hate speech but also notes that this should not inadvertently lead to removal of content which is not illegal.

The EU's Digital Services Act (DSA) is the latest attempt to combat unlawful expression on the Internet. Political agreement on the DSA was reached on 23 April 2022 between the European Parliament and EU Member States, and the final text of the DSA was published on 27 October 2022.[6] The DSA will enter into force in November 2022, but application of the provisions will begin mainly on 17 February 2024 (Article 93).

The DSA contains a common set of rules on responsibilities and accountability for providers of intermediary services and online platforms such as social media and marketplaces. It also aims at effective harmonisation of the legal framework in EU Member States, and provision of high levels of protection to all Internet service users by setting out notice-and-action procedures for illegal content, and the possibility to challenge platform content moderation decisions.[7] The DSA sets more stringent requirements and responsibilities for extremely large online platforms with more than 45 million European users.[8] Individual legislative measures at the national level would not be effective in ensuring uniform strong protection of rights online, so the regulation as such can be considered necessary.[9]

[2]Assemblée Nationale, 'PPL visant à lutter contre les contenus haineux sur internet: adoption en lecture définitive' (13 May 2019) <www.assemblee-nationale.fr/dyn/actualites-accueil-hub/ppl-visant-a-lutter-contre-les-contenus-haineux-sur-internet-adoption-en-lecture-definitive> accessed 4 October 2022.

[3]Bundesgesetz über Maßnahmen zum Schutz der Nutzer auf Kommunikationsplattformen (Kommunikationsplattformen-Gesetz – KoPl-G). StF: BGBl. I Nr. 151/2020 (NR: GP XXVII RV 463 AB 509 S. 69. BR: 10457 AB 10486 S. 917.)

[4]See <www.bills.parliament.uk/bills/3137> accessed 20 September 2022. Regarding regulating online harms, see J Woodhouse's text at House of Commons Library <www.commonslibrary.parliament.uk/research-briefings/cbp-8743/> accessed 25 April 2022. For a comprehensive summary on the topic with many references, see J Woodhouse, 'Regulating Online Harms' (2021) House of Commons Briefing Paper 8743, and regarding evaluation and criticism see e.g., V Nash, 'Revise and resubmit? Reviewing the 2019 Online Harms White Paper' (2019) 11 Journal of Media Law 18; L Price, 'Platform responsibility for online harms: towards a duty of case for online hazards' (2021) 13 Journal of Media Law 238.

[5]European Commission, 'Commission Recommendation of 1 March 2018 on measures to effectively tackle illegal content online',COM (2018) 1177 final.

[6]Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act).

[7]See DSA Chapter III Section 4.

[8]More on DSA Chapter IV Section 4.

[9]Lakivaliokunnan lausunto 9/2021, U 2/2021, 8 (statement of Finnish Legal Affairs Committee).

Additionally, non-legislative measures have been taken. On 30 June 2016, the Commission and four major IT companies (Facebook, Microsoft, Twitter and YouTube) presented a Code of Conduct on countering illegal hate speech online to ensure that requests to remove content are dealt with speedily. Since 2016, eight more IT companies have subscribed to the Code of Conduct.[10] Many platforms and actors have also set their own rules.

In current digital environments, the speed and wide reach of the Internet, the right to anonymous expression, and the difficulty of removing illegal content and enforcing judgments when offensive content has been made public from a country other than where the target of content lives, have led to steps being taken at both national and European levels to clarify and increase the responsibility of intermediary service and online platform providers. Indeed, it seems that imposing greater responsibility on them to remove unlawful content would be an effective way to prevent the spread of illegal content, because trials are lengthy and expensive and cannot even be initiated unless the right defendant – usually the writer of an offensive post – can be identified. However, the growing responsibilities and obligations of intermediary service providers and platform providers also involve risks to freedom of expression – for example, in the form of over-removal of content for fear of liability.[11] On the other hand, monitoring obligations may also interfere with the right to private life and personal data protection.[12] Therefore, it is important to map what kind of boundary conditions the European Convention of Human Rights (ECHR) sets for new regulation and present interpretation, which is why the practice of the European Court of Human Rights (ECtHR) is studied here.

Aim of the research

Member states of the Council of Europe have an obligation to guarantee human rights as described in the ECHR and interpreted in the practice of the ECtHR. This obligation also extends to relations between private parties.[13] The aim of this article is to identify and analyse the conditions under which administrators of online communication platforms can be held liable for content produced by others (users). The focus is on the criteria

[10]The EU Code of Conduct on Countering Illegal Hate Speech Online: The Robust Response Provided by the European Union < www.ec.europa.eu/info/policies/justice-and-fundamental-rights/combatting-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en#theeucodeofconduct > accessed 4 October 2022.

[11]F Erixon, 'Too big to care or "too big to share": The Digital Services Act and the consequences of reforming intermediary liability rules' (2021) European Centre for International Political Economy Policy Brief 5/2021, 1, 4, 8, 9.

[12]G Frosio and C Geiger, 'Taking Fundamental Rights Seriously in the Digital Services Act's Platform Liability Regime' (2022) European Law Journal (forthcoming).

[13]See, e.g., *Aksu v Turkey* [GC] App no 4149/04 and 41029/04 (ECtHR 15 March 2012) (59) and *Khurshid Mustafa and Tarzibachi v Sweden* App no 23883/06 (ECtHR, 16 December 2008) (30–5).

laid down by the ECtHR in its case law when defining the scope of liability of online content service providers. So, under what conditions can the administrator of an online communication forum be held liable for offensive content posted by a third party? This article also concludes how the new regulation fits within the boundary conditions set by the ECHR. This is important because, according to the Treaty of European Union (TFEU) Article 6(3), fundamental rights as guaranteed by the ECHR constitute general principles of Union law.[14]

This article focuses on administrators of online commenting platforms, not on intermediary or hosting services. The focus is on offensive – in terms of insulting and defamatory – expressions targeting a specific person. This text will not discuss ethnic agitation or other kinds of hate speech targeting a specific group of people. Nor will the discussion involve disseminating child pornography, expressions related to terrorism, or breaching intellectual property rights, all of which are covered by special regulation.

For example, Directive 2011/93/EU[15] requires that Member States take measures to remove web pages containing or disseminating child pornography and allows them to block access to those web pages. Directive (EU) 2017/541 contains similar provisions in respect of online content constituting public provocation to commit a terrorist offence.[16] Regulation (EU) 2021/784 on addressing dissemination of terrorist content online 'lays down uniform rules to address the misuse of hosting services for the dissemination to the public of terroristic content online.'[17] As for intellectual property rights, Directive 2004/48/EC on enforcement of intellectual property rights enables competent judicial authorities to issue injunctions against intermediaries whose services are being used by a third party to infringe an intellectual property right.[18] Moreover, the revised Audiovisual Media Services

[14]According to the Charter of Fundamental Rights of the European Union (CFR) (2000) C346/01, art 52(3), insofar as the 'Charter contains rights which correspond to rights guaranteed by' the ECHR 'the meaning and scope of those rights shall be the same as those laid down by the said Convention.' The CJEU has also applied and referred to the practice of the ECtHR (see, e.g., joined cases C-92/09 and C-93/09 Volker und Markus Schecke GbR and Hartmut Eifer v Land Hessen [2010] ECR I-11063 (51), (52), (72), (87); case C-400/10 PPU V.Mcb. v L.E. [2010] ECR I-08965 (53); C-237/15 PPU Minister for Justice and Equality v Francis Lanigan (56); C-562/13 Centre public d'action sociale d'Ottignies-Louvain-La-Neuve v Moussa Abdida (47); C-398/13 Inuit Tapiriit Kanatami and Others v Commission (45), (61).

[15]Directive 2011/93/EU of the European Parliament and of the Council of 13 December 2011 on combating the sexual abuse and sexual exploitation of children and child pornography and replacing Council Framework Decision 2004/68/JHA [2011] OJ L335/1.

[16]Directive 2017/541/EU 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] OJ L88/6.

[17]Regulation 2021/784/EU of the European Parliament and of the Council 29 April 2021 on addressing the dissemination of terrorist content online 2021/19/INIT [2021] OJ L 172.

[18]Directive 2004/48/EC of the European Parliament and of the Council of 29 April 2004 on the enforcement of intellectual property rights [2004] OJ L157/45.

Directive regulates illegal and harmful online content and extends application of the rules to certain social media platforms.[19]

Distinguishing hate speech (ethnic agitation and the like), terrorist information, and child pornography from offensive speech, the difference lies in that the first-mentioned categories are usually identifiable as illegal – at least to some extent – by reading the text or looking at a photo. Offensive language such as defamation or illegal dissemination of private information is more difficult to identify as illegal by just seeing the message(s). Defamation, for example, requires the absence of reasonable grounds for an offensive statement, and an online communication site administrator may simply be unable to check whether reasonable grounds exist for factual statements. In turn, the extent to which harsh language may be tolerated also depends on the status of the target and on their previous behaviour and language use. A service that runs an online discussion forum could be unable to judge whether statements are reasonably grounded or not. A duty to investigate the lawfulness or accuracy of comments could slow - and ultimately stop - debate.[20] So, a reasonable basis exists for treating different kinds of expressions differently in legislation. Related arguments also apply to disseminating private information, where the criteria laid down by the ECtHR involve a contribution to a debate of general interest, whether the person concerned is a public figure, the topic in question, and the form and consequences of publication. [21]

Freedom of expression and the right to private life on the internet

Freedom of expression is guaranteed in Article 11 of the Charter of Fundamental Rights of the European Union (CFR) and in Article 10 ECHR. The CJEU has held that Article 11 CFR is equivalent to Article 10 ECHR, including the case law of the European Court of Human Rights (ECtHR).[22]

[19]Directive 2010/13/EU of the European Parliament and of the Council of 10 March 2010 on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive) (2010) OJ L95/1, as amended by Directive 2018/1808/EU (2018) OJ L303/69). For a summary of the European regulatory framework see A Bertolini, F Episcopo and N Cherciu, 'Liability of online platforms' (2021) European Parliamentary Research Service, IV-IIIV.

[20]More on this, see Law Commission of Ontario, *Defamation Law in the Internet Age* (2020), 76–777, <www.lco-cdo.org/wp-content/uploads/2020/03/Defamation-Final-Report-Eng-FINAL-1.pdf> accessed 4 October 2022.

[21]These criteria were laid down in two Grand Chamber cases: *Axel Springer AG v Germany* (2012) 55 EHRR 6 (89–95), and *Von Hannover v Germany (No. 2)* (2012) 55 EHRR 15 (108–13), and are repeated in many judgments thereafter. For a more detailed view, see P Korpisaari, 'Balancing freedom of expression and the right to private life in the European Court of Human Rights – application and interpretation of the key criteria' (2017) 22 Communications Law Journal 39.

[22]See, e.g, Case C-163/10 Aldo Patriciello (Criminal Proceedings against) (2012) EU:C:2011:543 1 C.M.L.R. 11 [31]; or C-157/14 Société Neptune Distribution v Ministre de l'Économie et des Finances (2016) EU:C:2015:823 2 CMLR 24 [63–5].

According to Article 10 ECHR, freedom of expression includes the freedom to hold opinions as well as to receive and impart information and ideas. It constitutes 'one of the essential foundations of a democratic society and one of the basic conditions for its progress',[23] and it can be used to alert the public about matters of public interest.[24] Freedom of expression can also be regarded as an important value as such.[25] According to the ECtHR, freedom of expression also applies to expressions that offend, shock, or disturb.[26]

The Internet allows self-expression without the restrictions imposed by traditional media. The ECtHR has recognised the importance of the Internet for freedom of expression. Because of its accessibility and its capacity to store huge amounts of information, the Internet plays an important role in enhancing access to news and facilitating dissemination of information.[27] The Internet offers an unprecedented platform for spreading user-generated content,[28] and it is an important vehicle for citizen journalism, circulating ideas and information that the traditional media do not.[29] The ECtHR has also considered that blocking access to certain websites or online services can be a violation of Article 10[30] because access to the Internet provides essential 'tools for participation in activities and discussions concerning political issues and issues of general interest.'[31] Access and expression on the Internet may also attain social-inclusion and regional development goals in areas with deep economic and social inequalities.[32]

[23]European Court of Human Rights (ECtHR), e.g, *Barthold v Germany* (1985) 7 EHRR 383; *Handyside v the United Kingdom* (1979–80) 1 EHRR 737; *Zana v Turkey* (1999) 27 EHRR 667; *Von Hannover v Germany (No. 2)* (2012) 55 EHRR 15; *Axel Springer AG v Germany* (2012) 55 EHRR 6; *Gillberg v Sweden* App no. 41723/06 (ECtHR, 3 April 2012).

[24]See, e.g, *Heinisch v Germany* App no 28274/08 (ECtHR, 21 July 2011) and *Guja v Moldova* App no 14277/04 (ECtHR, 12 February 2008).

[25]T I Emerson, 'Toward a General Theory of the First Amendment' (1963) 72 The Yale Law Journal 877, 878–9; J Oster, *Media Freedom as a Fundamental Right* (CUP 2015), 13–20.

[26]E.g, *Hertel v Switzerland* (1999) 28 EHRR 534; *Steel and Morris v United Kingdom* (2005) 41 EHRR 22; *Stoll v Switzerland* (2008) 47 EHRR 59; *Mouvement raëlien suisse v Switzerland* (2013) 56 EHRR 14.

[27]*Delfi AS v Estonia* (2016) 62 EHRR 6 [133]; *Times Newspapers Ltd v the United Kingdom* (no. 1 and no. 2) App no 64367/14 (ECtHR, 13 Novemeber 2018) [27].

[28]*Delfi AS v Estonia* (2016) 62 EHRR 6 [110]; *Cengiz and Others v Turkey* App no 48226/10 and 14027/11 (ECtHR, 12 Decemeber 2015) [52].

[29]*Cengiz and Others v Turkey* [52]. ' … political content ignored by the traditional media is often shared via YouTube, thus fostering the emergence of citizen journalism. From that perspective, the Court accepts that YouTube is a unique platform on account of its characteristics, its accessibility and above all its potential impact, and that no alternatives were available to the applicants.'

[30]*Cengiz and Others v Turkey* [52]. The case considered YouTube, a video-hosting website on which users can upload, view and share videos; *Ahmet Yıldırım v Turkey* [49], regarding Google Sites, a Google service designed to facilitate the creation and sharing of websites within a group.

[31]*Cengiz and Others v. Turkey* (49).

[32]M Vazquez, M Rubio and A Morales, 'Thoughts on the Regulation of Content on Social Media in Latin America: Authors' Rights, Limitations, and Content Filtering' in B Holznagel, J Bauer, L Woods and P Korpisaari (eds), *Perspectives of Platform Regulation: Models and Limits* (NOMOS, 2021) 233–59.

The right to protection of private life covers a person's physical, psychological, and moral integrity,[33] as well as the right to establish and develop relationships with other human beings.[34] Activities of a professional or business nature may also belong to the concept of private life.[35] The right to a private life may also cover protecting reputation[36] and honour.[37] Article 8 does not protect against loss of reputation as the foreseeable consequence of one's own reprehensible action, such as commission of a criminal offence.[38]

Unfortunately, the potential created by social media is not always used for the common good. The speed of communication in social media and the ability to express oneself anonymously has increased the number of obscene insults towards individuals and ethnic, religious, and other groups.[39] As the ECtHR put it in *Delfi*, 'Defamatory and other types of clearly unlawful speech, including hate speech and speech inciting violence, can be disseminated like never before, worldwide, in a matter of seconds, and sometimes remain persistently available online.'[40]

Social media services such as Facebook and Twitter, media news commentary sites, and Internet discussion forums are common mediums for these attacks. The victim of a violation is in a vulnerable position since removing or correcting false information or identifying who sent a defamatory statement or photo can be difficult, even impossible.[41]

While social media services provide a useful platform for obtaining information, presenting and commenting on opinions, hate speech and insults on the Internet can also silence people who have been, or are afraid of becoming, targets. Hate speech can also lead to hostility, discrimination and even violence.[42]

[33] See, e.g, D Solove, *Understanding Privacy* (Harvard University Press 2008); J Marshall, *Personal Freedom through Human Rights Law?: Autonomy, Identity and Integrity Under the European Convention on Human Rights* (Martinus Nijhoff Publishers 2009); G González Fuster, *The Emergence of Personal Data Protection as a Fundamental Right of the EU* (Springer 2014).

[34] See B Roessler and D Mokrosinska (eds), *Social Dimensions of Privacy: Interdisciplinary Perspectives* (CUP 2015).

[35] *Niemietz v Germany* (1993) 16 EHRR 97; *Von Hannover v Germany (No. 2)* (2012) 55 EHRR 15. For an analysis, see P Korpisaari, 'Balancing freedom of expression and the right to private life in the European Court of Human Rights – application and interpretation of the key criteria' (2017) 22 Communications Law Journal 39.

[36] *Chauvy and Others v France* (2005) 41 EHRR 29; *White v Sweden* (2008) 46 EHRR 3; *Fürst-Pfeifer v Austria* App no 33677/10 and 52340/10 (ECtHR, 17 May 2016).

[37] *Radio France and Others v France* (2005) 40 EHRR 29; *A v Norway* App no. 28070/06 (ECtHR, 9 April 2009).

[38] *Sidabras and Džiautas v Lithuania* App no 55480/00 and 59330/00 (ECtHR, 27 July 2004) (49); *Axel Springer AG v Germany* (2012) 55 EHRR 6 (83). See also *Karakó v Hungary* App no 39311/05 (ECtHR, 28 April 2009) (22–3) and Oster's (2015) notion that the Court´s way of reasoning judgments in this respect has not always been consistent, 149–50.

[39] See K Koivukari and P Korpisaari, 'Online Shaming – a New Challenge for Criminal Justice' in Bayer et al 2021, 473–87, 475–9.

[40] *Delfi* (110).

[41] See the opinion of The Helsinki Foundation for Human Rights in *K.U. v Finland* App no 2872/02 (ECtHR, 2 December 2008) (34).

[42] For disadvantages of hate speech see e.g, U Kohl, 'Platform regulation of hate speech – a transatlantic speech compromise?' (2022) 14 Journal of Media Law 25.

Because this may affect both individuals and society at large, it is worth considering how to balance the competing rights to freedom of expression, on the one hand and the right to private life (including honour) on the other, in order to protect citizens from hateful and defamatory speech while mitigating the risks of erroneous or unjustified blocking or removal of comments.[43]

Website administrators: liability in the practice of the ECtHR

Margin of appreciation

The following offers an overview of website administrators' liability in the practice of the ECtHR. Importantly, the ECtHR establishes a European minimum level of protection. Therefore, it is not for the ECtHR to decide the case instead of the national courts as if it were an appeal instance, but to examine whether human rights have been violated. Since states always enjoy a margin of appreciation in their decision-making, the best possible outcome in a particular case cannot be concluded from ECtHR judgments. Indeed, in a Grand Chamber judgment of March 2016 the ECtHR repeated that in a conflict between Articles 10 and 8, '[w]here the national authorities have weighed up the interests at stake in compliance with the criteria laid down in the Court's case law, strong reasons are required if it is to substitute its view for that of the domestic courts.'[44] In practice, however, in several of its judgments the ECtHR has assessed cases in detail, which has also attracted criticism.[45]

Freedom of expression has traditionally enjoyed effective protection in the case law of the ECtHR, which has emphasised the essential role of the press as a "public watchdog". Additionally, NGOs have enjoyed wide protection of freedom of expression, and their role as important disseminators of information and opinions has been recognised.[46]

[43] See EC COM/2020/825, n 8, 12.

[44] *Bédat v Switzerland* App no 56925/08 (ECtHR, 29 March 2016) (54). This principle had been stated already in *MGN Limited v the United Kingdom* App no 39401/04 (ECtHR, 18 November 2011) (150) (see also (155)) and it is repeated in, e.g, *Axel Springer AG v Germany* (2012) 55 EHRR 6 (88), *von Hannover (no 2) v Germany* App no 40660/08 and 60641/08 (ECtHR, 7 February 2012) (107), *Palomo Sánchez and Others v Spain* App no 28955/06, 28957/06, 28959/06 and 28964/06 (ECtHR, 12 September 2011) (57), *Fürst-Pfeifer v Austria* App no. 33677/10 and 52340/10 (ECtHR, 17 May 2016) (40) and *Delfi AS v Estonia* App no 64569/09 (ECtHR, 16 June 2015) (139).

[45] J Kratochvil, 'The Inflation of the Margin of Appreciation by the European Court of Human Rights' (2011) 29 Netherlands Quarterly of Human Rights 324, 330, and K Hughes, 'Balancing Rights and the Margin of Appreciation: Article 10, Breach of Confidence and Success Fees' (2011) 3 Journal of Media Law 29, 40.

[46] *Animal Defenders International v the United Kingdom* App no 48876/08 (ECtHR, 22 April 2013) (103). See also *Youth Initiative for Human Rights v Serbia* App no 48135/06 (ECtHR, 25 June 2013) and *Magyar Helsinki Bizottság v Hungary* App no 18030/11 (ECtHR, 8 November 2016).

Liability exemption not applicable to administrators of online discussion forums

Regarding online discussion forums, a forum administrator does not typically familiarise themselves with content before it is published, nor do they make intentional decisions on publication. Here it differs from the "traditional media" (print publications, TV and radio programmes). In pre-moderated forums, the purpose of moderation is usually to eliminate unlawful expressions, or expressions unrelated to the topic of the forum – not to edit or moderate them or to check the background facts. Therefore, these forums cannot be directly compared to editorial journalistic content. Another distinction in comparison with the traditional media is that in the case of online discussion forums and user-generated content it might be difficult or impossible to identify the author of an expression, or even the person moderating the content. So, if the administrator of an online forum is not held responsible for published content, the result may be that in practice no one can be held liable.

Article 14 e-Commerce Directive establishes liability exemption for a hosting service provider when 'an information society service is provided that consists of the storage of information provided by a recipient of the service'.[47] According to Article 14, Member States must ensure exemption from liability in situations where the service provider does not have actual knowledge of illegal activity or information and, as regards claims for damages, is not aware of facts or circumstances from which illegal activity or information is apparent. Liability exemption also applies when the provider, upon obtaining such knowledge or awareness, acts promptly to remove or disable access to the information. Exemption from liability will not apply when the recipient of the service is acting under the authority or control of the provider.

Article 15 further stipulates that a general obligation shall not be imposed on providers to monitor the information they transmit or store. Provision of hosting services covered by the liability exemption applies only to passive activities. This means that the administrator of an online discussion forum who sets the rules of the forum and can moderate content according to its own decisions is not excluded from liability according to this article. However, the liability exemption can be enjoyed by a service provider who offers storage space for the forum. For example, in *Eva Glawischnig-Piesczek* the European Court of Justice (CJEU) considered that Facebook Ireland Limited was providing hosting services via its global social media platform.[48] The purpose of the provision is that operators who are not able to affect the content of online information transmitted or stored are not held responsible for it.

[47]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) [2000] OJ L178/1.

[48]C-18/18 Glawischnig-Piesczek v Facebook Ireland Limited [2016] ECLI:EU:C:2019:821.

The CJEU held in *Google France SARL and Google Inc. v Louis Vuitton Malletier SA*[49] and *Google France SARL v Viaticum SA and Luteciel SARL*[50] that an Internet referencing service provider was covered by the liability exemption if it had not played an active role of such a kind as to give it knowledge of, or control over, the data stored. In the *Frank Peterson* and *Elsevier* cases[51] the CJEU held that activities by the operator of a video-sharing platform or a file-hosting and file-sharing platform fell within the scope of the liability exemption provided that the operator did not 'play an active role of such a kind as to give it knowledge of or control over the content uploaded to its platform.'

In *L'Oréal SA and Others v eBay International AG and Others*[52] the CJEU considered that optimising the presentation of offers or promoting them constituted an active role, which fell outside the scope of the liability exemption. In *Sotiris Papasavvas v O Fileleftheros Dimosia Etairia Ltd, Takis Kounnafi and Giorgos Sertis*[53] the CJEU held that the limitations of liability listed in Articles 12–14 did not apply to a newspaper publishing company which operated a website on which the online version of a newspaper was posted and which was remunerated through income generated by commercial advertisements posted on that website, since it had knowledge of the information posted and exercised control over that information.

The distinction between passive or active as a criterion for liability has been criticised, taking into consideration the extensive moderation (filtering, sorting and optimising content) that platforms undertake as part of their business model.[54] However, it should be noted that the fact that the liability exemption as such is not applicable does not itself constitute liability. The basis for liability must always be found in legislation or court practice.

Liability exceptional in ECtHR case law

Landmark 'Delfi' case: liability for 320 euros compensation not a violation of freedom of expression

The judgment in *Delfi AS v Estonia*[55] is examined here in more detail as it is a judgment of the Grand Chamber and the first judgment in which the ECtHR assessed whether online discussion platform administrators could be held

[49]C-236/08 Google France SARL and Google Inc v Louis Vuitton Malletier [2010] ECLI:EU:C:2010:159.
[50]C-237/08 Google France SARL v Viaticum SA and Luteciel SARL [2010] ECLI:EU:C:2008:389.
[51]C-682/18 Peterson v Google LLC, YouTube LLC, YouTube Inc and Google Germany GmbH and C-683/18 Elsevier Inc v Cyando AG [2020] ECLI:EU:C:2020:586.
[52]C-324/09 L'Oréal SA and Others v eBay International AG and Others [2011] ECLI:EU:C:2011:474.
[53]C-291/13 Sotiris Papasavvas v O Fileleftheros Dimosia Etairia Ltd, Takis Kounnafi and Giorgos Sertis, (2014) ECLI:EU:C:2014:2209-
[54]M C Buiten, 'The Digital Services Act: From Intermediary Liability to Platform Regulation' (2021) Working Paper available at SSRN 3876328, accessed 15 July 2021, 4, 15, 17.
[55]*Delfi AS v Estonia* (2016) 62 EHRR 6.

liable for content posted on their site by third parties. The ECtHR itself has mirrored its subsequent judgments against this so-called landmark case.

Delfi, a large professionally managed commercial news portal, published an article on the destruction of ice roads by a ferry company and within a couple of days 185 comments were made on the news-related discussion, of which about twenty contained threats and offensive language targeting the director of the ferry company. The offensive comments were removed six weeks after they were uploaded on the website, immediately after Delfi had received notification of the content from the lawyer acting for the individual targeted by these comments.

Delfi's website contained rules stating that the authors of comments were responsible for the content they published. Furthermore, Delfi carried out moderation of illegal content by automatically deleting certain words and by a notification system that allowed the reader to report illegal content. Despite these protective measures, the target of the comments was awarded EUR 320 in compensation for non-pecuniary damage. Delfi appealed to the ECtHR claiming that its freedom of expression had been violated.

The ECtHR considered that the comments were mainly hate speech and speech inciting violence towards the director of the ferry company.[56] Thus, their illegality was obvious and did not require any linguistic or legal interpretation. It was therefore not freedom of expression of the authors that was to be assessed but only whether the decision of the national court infringed Delfi's right to disseminate information and opinions.

The ECtHR noted that Delfi ran on a commercial basis and that due to the nature of the Internet the obligations and responsibilities of an Internet news portal might differ from the liability of traditional media publishers for third-party content. Even though the right to publish anonymously on the Internet was regarded as an important value, it had to be balanced against other rights and interests. Because of the anonymity of the comments, if was difficult to identify their authors. Shifting the risk of liability for damages from the defamed person to the media company, which was usually in a better financial position than the defamer (if one were even found), was not such a disproportionate interference with the media company's right to freedom of expression as to constitute an infringement of freedom of expression. On that basis, the ECtHR held, by fifteen votes to two, that the

[56]The comments are cited in para 18 of the judgment. The first five are: '1. (1) there are currents in [V]äinameri; (2) open water is closer to the places you referred to, and the ice is thinner. Proposal – let's do the same as in 1905, let's go to [K]uressaare with sticks and put [L.] and [Le.] in a bag; 2. bloody shitheads … they're loaded anyway thanks to that monopoly and State subsidies and have now started to worry that cars may drive to the islands for a couple of days without anything filling their purses. burn in your own ship, sick Jew!; 3. good that [La.'s] initiative has not broken down the lines of the web flamers. go ahead, guys, [L.] into the oven!; 4. [little L.] go and drown yourself; 5. aha … [I] hardly believe that that happened by accident … assholes fck'.

judgment concerning the claims for damages did not infringe freedom of expression. The Chamber had also reached the same conclusion.

When examining the case, the explicit limitation constituted by the ECtHR must be noted: it stated that the case was not about a forum where third-party comments on any topic could be disseminated without input from the forum administrator as to content creation. Nor did the case concern a social media platform where the platform provider does not itself offer any content and where the content provider might instead be a private person running the website or blog as a hobby. In this way, the ECtHR drew a distinction between an active website administrator and a passive one, as well as between commercial and non-commercial activities. The ECtHR also paid attention to the context of the comments, the measures taken by the online discussion provider to prevent publication of or remove offensive comments, the possibility of holding the authors of the comments liable, and the penalties imposed on the platform provider. In addition, the ECtHR indicated that Delfi allowed unregistered users to comment and did not exercise stricter control over comments, although in this case, in the view of the ECtHR, the risk of negative comments was higher than usual.[57]

The case has been heavily criticised. The dissenting judges of the ECtHR, Sajó and Tsotsoria, considered that the position of the majority of the judges would mean that, in the future, comments created by third parties would have to be monitored from the moment they are posted, as a result of which comments features may no longer be offered or comments might be removed too easily.[58] Since Delfi had taken several measures to prevent or remove illegal content, the concern of the minority was easy to understand. On the other hand, the threats remained on the page for six weeks without being noticed by the moderators or the automated editorial system, even though, according to the ECtHR, the content was clearly unlawful hate speech and thus there was no uncertainty as to their illegality. Moreover, this was not a criminal judgment but a judgment regarding 320 euros in compensation. In addition, the state enjoys a certain margin of appreciation.

The judgment has also been criticised on the grounds that the ECtHR based its argumentation on its own practice without taking into account, for example, the liability regime established in the e-Commerce Directive, or other international principles of interpretation relating to the liability of intermediaries.[59] There have also been concerns that, due to the *Delfi* case, the CJEU will reduce the number of operators covered by exemptions

[57] See J Barata Mir and M Bassini, 'Freedom of Expression in the Internet: main trends of the case law of the European Court of Human Rights' in O Pollicino and G Romeo (eds), *The Internet and Constitutional Law: The Protection of Fundamental Rights and Constitutional Adjudication in Europe* (Routledge 2016), 71, 84.

[58] Similarly, L Brunner, 'The Liability of an Online Intermediary for Third Party Content' (2016) 16 Human Rights Law Review 163, 172.

[59] Brunner (2016) (n 58) 167–9, and Barata Mir and Bassini (2016) (n 57) 89–92.

from liability set out in the e-Commerce Directive.[60] However, it is clear that Delfi could not enjoy the protection afforded to hosting service providers, because it did not merely provide passive storage space but published and sought to attract comments on news, set out the 'rules' for the comments section, and was entitled to moderate the content of comments in accordance with its own principles.

One criticism is that the ECtHR assumes that content reported is always illegal and that reporters act in good faith, or that different filtering programmes would provide an easy solution to the issue.[61] It has been considered that the judgment imposes an additional obligation on Internet news portals struggling with profitability as they carry out their role as mediators of public debate.[62] Comment sections could lose their nature as free channels for sharing data and information and self-expression if their content were to become "accepted" by columnists.[63] However, the ECtHR emphasised in this case that the unlawfulness of the content was obvious, and that evaluation has to be case by case.

In addition, the case concerned civil liability for damages and not criminal liability. In modern tort law, damages do not necessarily include the same assessment of reprehensibility as in traditional tort law, because society and technological developments have created new risks, the occurrence of which is difficult to predict in some cases.[64] Liability of business companies is increasingly based on operating in a certain position of responsibility,[65] and damages have begun to be seen as part of a company's business risk and the costs it has to pay, whether the damage was caused by reprehensible conduct or not.[66] Online platform administrators are often in the best position to manage risks and ensure that compensation is paid.[67]

[60] Brunner (2016) (n 58) 173–4. See also D Voorhoof, 'Delfi AS v. Estonia: Grand Chamber confirms liability of online news portal for offensive comments posted by its readers' (Strasbourg Observers, 18 June 2015) <www.strasbourgobservers.com/2015/06/18/delfi-as-v-estonia-grand-chamber-confirms-liability-of-online-news-portal-for-offensive-comments-posted-by-its-readers> accessed 3 October 2022.

[61] See, e.g, Art 19, 'European Court strikes serious blow to free speech online' (14 October 2013) <www.article19.org/resources/european-court-strikes-serious-blow-free-speech-online> accessed 3 October 2022.

[62] Brunner (2016) (n 58) 164.

[63] M Maroni, 'A Court's Gotta Do, What a Court's Gotta Do. An Analysis of the European Court of Human Rights and the Liability of Internet Intermediaries through Systems Theory' (2019) RSCAS 2019/20 EUI Working Papers 1, 2; M Maroni, 'The liability of internet intermediaries and the European Court of Human Rights' in B Petkova and T Ojanen (eds), *Fundamental Rights Online, The Future Regulation of Internet Intermediaries* (Edward Elgar 2020), 258.

[64] T Wilhelmsson, *Senmodern ansvarrsrätt. Privaträtt som redskap för mikropolitik* (Kauppakaari 2001), 203.

[65] M Hemmo, *Sopimus ja delikti. Tutkimus vahingonkorvausoikeuden vastuumuodoista* (Lakimiesliiton kustannus 1998).

[66] E Hoppu, 'Vahingonkorvauksen kehityslinjoja', (1998) 6–7 *Lakimies* 1048, 1051. See also, T Wilhelmsson, 'Vastuu ja yksityisoikeuden systeemi' (1997) 8 *Lakimies* 1180, 1200.

[67] A Bertolini, F Episcopo and N Cherciu, 'Liability of online platforms' (2021) European Parliamentary Research Service, IX.

However, the deterrent effect on exercise of freedom of expression must always be taken into account when considering liability for damages. It is not only the quantum of damages that matters, but in some cases the mere imposition of a symbolic amount may infringe upon freedom of expression by including a statement of principle that the exercise of freedom of expression was unlawful.[68]

No post-'Delfi' liability until 'Sanchez' in 2021

For a long period after *Delfi*, platform administrators were not held liable. In *Włodzimierz Kucharczyk v Poland*,[69] delivered five months after *Delfi* and which has not attracted much attention, a lawyer's name was listed on a privately-run Internet portal together with some 800 other lawyers. Of eighteen comments posted anonymously in three-and-a-half years and that concerned his professional skills, fifteen comments were highly favourable to him. One comment was negative, advising not to use his services since he was claimed to be 'utterly ignorant of his job' and 'disorganised and incompetent'. The administrator of this online forum refused to remove the post and the police discontinued their investigation regarding alleged defamation because the perpetrator had not been identified. Additionally, the lawyer's civil claim for removal of the content and a ban on future comments was dismissed, since in his capacity as a lawyer performing a public service, he had to accept public assessment and opinion. The comments were also not considered particularly defamatory. The lawyer complained to the ECtHR.

The ECtHR repeated the oft-mentioned legal homily: 'In order for Article 8 to come into play, an attack on a person's reputation must attain a certain level of gravity and in a manner causing prejudice to personal enjoyment of the right to respect for private life.'[70] The ECtHR considered that although an attorney had 'a different status to that of a judge or a prosecutor', his profession was 'nevertheless one of public trust' and 'comments on a lawyer's professional skills' constituted 'matters in which the community at large had an interest'. The statement was also an opinion and a value judgment, not a statement of fact or an allegation of unlawful or improper conduct on the part of the applicant. The critical comment was also followed by many positive comments assessing him as a very good lawyer. Considering the margin of appreciation and the circumstances, the ECTHR declared the application inadmissible.[71]

[68]E.g, *Brasilier v France* App No 71343/01 (ECtHR, 11 April 2006).
[69]App no 72966/13 (ECtHR, 17 December 2015).
[70]See *A v Norway* App no. 28070/06 (ECtHR, 9 April 2009) [64].
[71]See also a recent judgment of the Norwegian Supreme Court (Norges Høyesterett) 7 December 2021 (HR-2021-2403-A, (sak nr 21-055809SIV-HRET), where a website ('Legelisten') published opinions by doctors' patients about the doctors they had consulted. The applicant, who was a doctor, referred to GDPR 6(1)(f) claiming that there was no legitimate interest in publishing information about him.

The outcome of the decision can be considered correct. The statement did not amount to hate speech but, rather, was an opinion or a value judgment. However, one could disagree with the ECtHR's statement that 'the applicant, as a practising lawyer, should have accepted that he might be subjected to evaluation by anyone with whom he had ever had any professional dealings' for the reason that because the portal did not require users to register their personal information, and the technical design of the server did not allow for the storage of users' IP addresses, it was not possible to confirm whether the lawyer's critic had ever been his client. But this anonymity also reduces the credibility of the opinion.

The case of *Magyar Tartalomszolgáltatók Egyesülete and Index.hu Zrt v Hungary*[72] was decided soon after *Delfi*. Here the appellants were a non-profit self-regulatory body for the Internet industry and a news portal. A reader's opinion had been published on their website. The opinion rightly criticised a company for its misleading business practices. However, the comments were presented in a vulgar way.[73] The online platform administrators, held liable for these comments by the national courts, complained about a violation of their freedom of expression.[74] The website had posted a disclaimer of author's liability as well as a system that removed messages reported to contain illegal content.

According to the ECtHR, the comments were vulgar but not clearly illegal. A legal person could not be the object of an infringement of personal rights, and the reputation of a commercial company targeted by the comments did not enjoy protection similar to that of an individual person. In addition, there was general interest regarding the practices of the company criticised, and the subject matter concerned many consumers and Internet users. The comments had been deleted as soon as they had been reported – although this happened after the proceedings had started.

The ECtHR held that if website administrators were held liable, this could form the basis for subsequent claims in damages. In addition, liability would have a negative impact on the commenting environment in online discussion forums. The difference between this case and *Delfi* was that the messages were tasteless but not clearly illegal or hate speech. The topic of questionable business practices by companies was a matter of general interest and the discussion included important consumer information. The outcome of the case is consistent with the ECtHR's previous practice, in which issues of public interest have enjoyed greater protection of freedom of expression.

In this case, which balanced freedom of expression with the right to private life, the courts at all instances decided in favour of patients' freedom of expression.

[72] *Magyar Tartalomszolgáltatók Egyesülete and Index.hu Zrt v Hungary* App No 22947/13 (ECtHR, 2 February 2016).

[73] One comment stated that 'People like this should go and shit a hedgehog and spend all their money on their mother's tombs until they drop dead'.

[74] The appellants were ordered to pay legal expenses.

In *Pihl v Sweden*[75] the blog of a non-profit association alleged that a person named Pihl was involved in a Nazi party. The site stated that comments are not moderated in advance and that comments must comply with the law and good manners. An anonymous person commented on the blog post and said that he had heard from several people that Pihl was a hashish user.[76] Pihl commented on the blog, asserting that the claim was false and demanding its removal. Nine days later, the association deleted the blog post with its comments and published an apology. Pihl filed a civil lawsuit against the association, claiming symbolic damages of one crown (EUR 0.10) for defamation. When the claim was rejected, Pihl appealed to the ECtHR, arguing that the authorities had not safeguarded his reputation. The ECtHR dismissed the claim as unfounded. While the comment had been offensive, it amounted to neither hate speech nor an incitement to violence. The comment had been published in a small blog maintained by a non-profit association, where it had been viewable for only nine days and removed immediately the day following the applicant's complaint. Pihl had also found out the commenter's IP address, but it was not clear whether he had taken action to establish the commenter's identity to make him accountable for his actions. According to the ECtHR, the national authorities had reached the right solution by dismissing the complaint asserting the association's liability.

The ECtHR's ruling has been criticised for placing too much emphasis on the fact that the association hosting the blog was small and non-commercial. Indeed, it has been pointed out that this could lead to a further increase in the responsibility of bigger operators.[77] The question whether the performance of large commercial websites is assessed differently from that of small non-commercial websites is relevant. Freedom of expression also covers commercial activities, where an operator's advantageous financial position enables them to acquire monitoring tools and site maintenance staff.

In the case of *Tamiz v United Kingdom,*[78] the applicant was a local politician who had used degrading language towards women, for example calling local women 'whores'. He was criticised in a blog called 'London Muslim', as well as in its comment section, where he was described by numerous offensive expressions, such as asserting him to be a well-known drug dealer and a violent racist. Even his entire family was claimed to be criminal. The applicant complained about the offensive comments to Google Inc., which maintained the blog on its blogger.com platform. After a few

[75] *Pihl v Sweden* App no 74742/14 (ECtHR, 9 March 2017).
[76] 'that guy pihl is also a real hash-junkie according to several people I have spoken to.'
[77] D Voorhoof, 'Pihl v. Sweden: non-profit blog operator is not liable for defamatory users' comments in case of prompt removal upon notice' (Strasbourg Observers, 20 March 2017) <www.strasbourgobservers.com/2017/03/20/pihl-v-sweden-non-profit-blog-operator-is-not-liable-for-defamatory-users-comments-in-case-of-prompt-removal-upon-notice> accessed 4 October 2022.
[78] *Tamiz v United Kingdom* App no 3877/14 (ECtHR, 19 September 2017).

months, the applicant filed a defamation lawsuit against Google Inc., after which Google Inc., with the applicant's permission, sent the information to the blog administrator. The blog post with its comments was thereafter removed by the blog administrator. The applicant's defamation lawsuit was dismissed, and he thus brought his complaint to the ECtHR.

The ECtHR reiterated its previously stated principle that defamation must meet a certain threshold of seriousness in order to qualify as a human rights violation. As a politician, the applicant had to tolerate vulgar comments, which readers did not take seriously, according to the ECtHR. The Court agreed with the national courts that while the majority of comments about which the applicant complained were undoubtedly offensive, they were mostly little more than 'vulgar abuse', which was common in communication on many Internet portals. The blog post and its comments were removed after Google Inc.´s notification. The ECtHR declared the applicant's claim 'manifestly ill-founded' and inadmissible.

Tamiz differed from *Pihl* in that the claims were directed against the service provider on whose platform the online forum was located, instead of against the online discussion forum administrator. This is probably because the identity of the administrator or owner of the blog could not be found. What both cases had in common was that the offensive messages were removed, thus reducing the damage caused to the applicants.[79]

In *Høiness v Norway*,[80] concerning the liability of a website administrator, the applicant was a well-known Norwegian lawyer and former TV personality. The online newspaper talked about her relationship to an elderly widow she would later inherit from and also opened a conversation on the topic, where vulgar and sexist comments were made about the applicant. Høiness had claimed approximately EUR 25,000 as compensation for damages from the website administrator and approximately EUR 2,500 from its supplier for defamation. The national court did not consider the comments offensive, and ordered the applicant to pay the legal expenses of the other parties (approximately EUR 20,000). The applicant complained to the ECtHR about violation of her private life.

The ECtHR did not consider the comments to be hate speech and therefore did not deem it necessary to examine them in depth. It would have been very difficult for the applicant to make claims against anonymous commenters. The news portal was large, and its popular forums were an extension of published articles. The content was moderated, and the website had a button for anyone to report illegal content. One could also send removal requests via email. Two offensive comments had been removed some thirteen minutes after the remark, and a third comment on the moderator's own initiative.

[79]Cf *Einarsson v Iceland* (2018) 67 EHRR 6 (vote 5–2); *Einarsson v Iceland (No. 2)* (2020) 70 EHRR 3.
[80]*Høiness v Norway* (2019) 69 EHRR 19.

According to the ECtHR, the measures taken by the media company and its editors to moderate the comments had been appropriate and the news portal could not be held responsible for comments submitted by third parties. There was therefore no infringement of Article 8 ECHR. Nor were the legal costs so high that Article 8 ECHR could have been infringed. The case demonstrates once again that expeditious removal of offending material is of great importance when assessing liability.

In the latest case of *Sanchez v France*[81] the Court held, by six votes to one, that the conviction of a politician for failing to promptly delete unlawful comments posted by third parties on his public Facebook wall, which was managed by him personally, did not amount to violation of his right to freedom of expression. This case was referred to the Grand Chamber on 17 January 2022 and the Grand Chamber held a hearing on 29 June 2022.

Mr Sanchez, a local councillor standing for election to Parliament at that time, had published a post criticising FP, his political opponent and a member of the European Parliament (MEP) and first deputy to the mayor of Nîmes. Two persons commented on the post with insulting expressions against Muslims, associating them with crime and insecurity by equating them with 'drug dealers and prostitutes' who 'reign supreme', 'scum who sell drugs all day long' or those responsible for 'throwing stones at white people's cars.' One comment also referred to FP's partner LT.

When LT became aware of the comments she felt directly and personally insulted about these racist comments and asked one of the commentators to remove his comment, which he also did immediately after her request. Then Mr Sanchez asked users to monitor the content of their comments but did not remove any of the comments already posted. Later the criminal court found him and the two commentators guilty of inciting hatred or violence against a group of people on the grounds of their origin or their membership or non-membership of a specific ethnic group. Mr Sanchez was ordered to pay a fine of 3 000 euros and 1 000 euros as compensation to LT.

The ECtHR pointed out that the comments were clearly unlawful and targeted at a specific group of people. According to the ECtHR, the foundation of a democratic and pluralistic society lies in tolerance and respect for the equal dignity of all human beings. Freedom of expression is wide in the electoral context, but politicians bear a particular responsibility in combating hate speech. The Court reasoned that Mr Sanchez had undertaken a duty to monitor the content published on his Facebook wall when he decided to make his wall public and allow his friends to post comments there. In addition, his political status required even greater vigilance on his part. Despite this, some comments had been visible for six weeks. The ECtHR

[81] *Sanchez v France* App no 45581/15 (ECtHR, 2 September 2021).

adjudged that, having regard to the margin of appreciation, there had been relevant and sufficient reasons for Mr. Sanchez's conviction.

The outcome of the chamber has been criticised for not being consistent with EU law (particularly Article 15 of the e-Commerce directive) and ECtHR case law and not taking into account the political context where the discussion took place.[82] However, maintaining a Facebook site in onés own name and opening and participating in a debate there are not activities that fall within the scope of liability exemption because these activities are not hosting services referred to in Article 14 of the e-Commerce Directive.

Unlike *Delfi*, this was a matter of criminal liability, for which the conditions are stricter than those for compensation liability. However, it is worth noting that international conventions also require criminalising acts of a racist and xenophobic nature, which are considered to constitute a violation of human rights and a threat to the rule of law and democratic stability.[83]

Conclusions: the current state of play

The ECtHR has understood the special characteristics of the Internet and its online discussion forums. Communication is fast, global, but sometimes thoughtless or even reckless. It can be difficult or impossible to hold the authors of comments liable.[84] On the other hand, due to the nature of online discussion forums, users do not usually read messages with the same seriousness as they would with journalistic content.[85]

In sum, the liability exemption of the e-Commerce Directive is not applicable to Facebook accounts, onés own blogs, discussion forums or other websites where the holder of that site creates content on their own or/and can define the rules or topics for discussion and has the power to moderate it.

So far, liability has arisen only in *Delfi* and *Sanchez*, where the expressions were clearly unlawful hate speech that targeted a member of a certain

[82]'Case Law, Strasbourg: Sanchez v France, Politician fined for failing to delete Facebook hate speech, no violation of Article 10' (Inforrm, 6 October 2021) <www.inforrm.org/2021/10/06/case-law-strasbourg-sanchez-v-france-politician-fined-for-failing-to-delete-facebook-hate-speech-no-violation-of-article-10/> accessed 4 October 2022.

[83]Additional Protocol to the Convention on Cybercrime, concerning the criminalisation of acts of a racist and xenophobic nature committed through computer systems, ETS 189, Strasbourg, 28.I.2003. See also ACTS ADOPTED UNDER TITLE VI OF THE EU TREATY COUNCIL FRAMEWORK DECISION 2008/913/JHA of 28 November 2008 on combating certain forms and expressions of racism and xenophobia by means of criminal law.

[84]*Delfi AS v Estonia* (2016) 62 EHRR 6 (142–3).

[85]See *Magyar Tartalomszolgáltatók Egyesülete and Index.hu Zrt v Hungary (*n 72) (77). 'Without losing sight of the effects of defamation on the Internet, especially given the ease, scope and speed of dissemination of information (see *Delfi AS*, cited above, (147)), the Court also considers that regard must be had to the specifics of the style of communication on certain Internet portals. For the Court, the expressions used in the comments, albeit belonging to a low register of style, are common in communication on many Internet portals – a consideration that reduces the impact that can be attributed to those expressions.'

minority, although in *Delfi* the hate speech might have originated rather because of the target´s business actions than his Jewish origin. In both cases, the unlawfulness of comments was abundantly clear without further examination.

In *Delfi,* the comments had been removed immediately after notification. In *Sanchez,* the comments were on his publicly accessible Facebook account, which he managed personally but did not take prompt action to delete comments posted by others. According to the Additional Protocol to the Convention on Cybercrime (ETS 189) member countries shall criminalise distributing or otherwise making available racist and xenophobic material to the public through a computer system when the acts are committed intentionally and without the right to do so. This means that the Convention allows – or even requires – criminalising failure to remove such content if the person who could remove content has actual knowledge about it.

No general monitoring requirement applies to Facebook accounts, blogs, Youtube channels, and so on, but when someone is notified that their website contains illegal material, this is strong evidence that they have actual knowledge about it. As mentioned earlier, not all crimes are punishable due to omission to react or prevent the consequences of a crime. However, ethnic agitation is an exception in this regard.

Additionally, a distinction must be drawn between civil and criminal liability. The conditions for criminal liability are more stringent than those for civil liability. When a website is held by a commercial business company or otherwise for commercial purposes, more actions can be required than when a website is run as a hobby.[86]

So far, the liability of online discussion forum administrators has arisen only exceptionally. Quick removal of defamatory or privacy-infringing material and measures applied by the applicant company to prevent or remove defamatory comments have usually had an exculpatory effect. Other relevant aspects are the context of comments, the liability of the actual authors of comments as an alternative to the intermediary's liability, and the consequences of domestic proceedings for the applicant company.[87] From a victim's point of view, the most crucial factor is often that insulting/offending content is removed from the public eye.

Moreover, other means are available to reduce harm and damage caused by defamatory or privacy-infringing material on the web. One can plead the right to be forgotten, which is a right guaranteed by Article 17 of the GDPR. For example, in the well-known *Google Spain* case the CJEU ruled that Google, as a commercial search firm that gathered personal information

[86]In addition to the case law presented here, see the European Court of Human Rights Press Unit, 'Fact-sheet – Hate speech' (2019) <www.echr.coe.int/Documents/FS_Hate_speech_ENG.pdf> accessed 4 October 2022.

[87]*Delfi AS v Estonia* (2016) 62 EHRR 6 (142–3).

for profit, had to remove links to private information when asked, provided that the information was no longer relevant.[88]

Some reflections on the DSA

Introduction and applicability of the DSA

Currently no uniform standard has been set for the liability of online platform administrators in Europe. Member states of the ECHR enjoy a certain margin of appreciation. Protection of privacy may, in some cases, even entail an obligation to uphold legislation that allows balancing between privacy and freedom of expression, as well as holding the perpetrator criminally liable or the online discussion platform administrator accountable for civil liability.[89] In addition, equality and non-discrimination of all persons require certain actions against misuse of freedom of expression. Because the DSA does not define what kind of content is legal or illegal, or the conditions for civil or criminal liability, it probably does not create a uniform liability regime.

The DSA is applicable to 'intermediary services offered to recipients of the service that have their place of establishment or are located in the Union, irrespective of where the providers of those intermediary services have their place of establishment.' (Article 2(1)). The particular scope of application is intermediary services consisting of services known as 'mere conduit', 'caching' and 'hosting' services (Article 1(2), 2(1-2) and Article 3 ((g)(i – iii)). This means that the DSA is not applicable to persons or entities that run a blog or web discussion forum or discuss on their own Facebook account or other platforms that create content themselves or together with users or that are set up for the purpose of publishing user-generated content.[90]

However, this regulation is also important for platform administrators because, if they fail to remove content that has been claimed as unlawful, removal can be asked from the intermediary service provider. For example, the DSA could have been applicable in the case of *Tamiz*, where Google Inc. as hosting services provider was the defendant instead of the blog administrator. Of course, applicability does not yet mean that content should have been removed. Thus, as to preventing dissemination of illegal

[88]Case C-131/12 Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González (Google Spain) (2014) EU:C:2014:317 EUECJ. See also ECtHR case *Hurbain v Belgium* App no 57292/16 (ECtHR, 22 June 2021) and compare to *ML and WW v Germany* App no 60798/10 and 65599/10 (ECtHR, 28 June 2018).

[89]See, e.g, *KU v Finland* (2009) 48 EHRR 52 and *Von Hannover v Germany (No. 1)* (2005), 40 EHRR 1.

[90]According to DSA 2(2) it is not applicable 'to any service that is not an intermediary service or to any requirements imposed in respect of such a service, irrespective of whether the service is provided through the use of an intermediary service, irrespective of whether the service is provided through the use of an intermediary service.'

content, the DSA does not target those closest to publishing the content (platform administrators), but primarily targets technical actors who enable the use of Internet-based services and who are passive in relation to published content. This is a relevant factor when considering how far-reaching intervention obligations can be imposed.

The transnational nature of situations causes problems, as messages can be published from servers in other countries and are visible in many countries.[91] Furthermore, it is easy to republish content which has been ordered to be deleted.[92]

Notice and take down mechanism and other new duties

Hosting service providers have been assigned the power to judge what content should be removed and what content can stay. From the victim´s standpoint the good thing is that, compared to court proceedings, Internet service providers can operate quickly and flexibly, and the procedure is usually cost-effective or free. It is favourable for the subject of an infringement if illegal content can be removed rapidly. How rapidly service providers must react depends on the facts, circumstances, and types of illegal content involved.[93] One more advantage for the victim of unwanted publicity is that this process usually attracts less public attention than a trial.

Service providers must fulfil some obligations to promote transparency. All providers of intermediary services must provide points of contact (Articles 11 and 12) and 'information on any policies, procedures, measures and tools used for the purpose of content moderation, including algorithmic decision-making, and human review, as well as the rules of procedure of their internal complaint handling system' (Article 14(1)). Very large online platforms must produce a transparency report (Article 15), which according to Article 15(1)(b) must include information on how many notices they have received and what kind of alleged illegal content these notices are about and 'the number of notices processed by using automated means and the median time needed for taking the action.'

According to Article 16, hosting services must maintain a notice and action mechanism where any individual or entity can notify illegal content. The request must contain a sufficiently substantiated explanation of the reasons why the content is illegal, and the decision must be made in a timely, diligent, non-arbitrary and objective manner. Hosting service providers are allowed to use automated means for processing and decision-making, but in that case they must provide information on such use.

[91]See cases C-251/20 Gtflix TVv DR (2021) EU:C:2021:1036; C-194/16 Bolagsupplysningen and Ilsjan (2017) EU:C:2017:766; C-509/09 and C-161/10 eDate Advertising and Others (2011) EU:C:2011:685.
[92]However, see C-18/18 Glawischnig-Piesczek v Facebook Ireland Limited (n 48).
[93]DSA rec 87.

Automated decision-making raises certain concerns because it can be too simplistic and rough, as well as leading to removal of legal content. On the other hand, entities that produce a large amount of illegal content can come up with ways to present certain expressions so that AI does not recognise them as illegal. Considering the large volume of unlawful content and the fact that a large part of it is clearly illegal, the introduction of an automatic decision-making system is probably quite necessary in the operations of larger operators.

Reasons for removing or blocking access to content must be explained (Article 17). Providers of very large online platforms must also operate an internal complaint-handling system free of charge. Decisions must be made under the supervision of appropriately qualified staff, and not solely on the basis of automated means.

Problems in evaluating unlawfulness of content and other vague points

Article 3(h) of the DSA defines that '"illegal content" means any information that, in itself or in relation to an activity, including the sale of products or the provision of services, is not in compliance with Union law or the law of any Member State which is in compliance with Union law, irrespective of the precise subject matter or nature of that law.' Recital 12 stipulates that:

> the concept of 'illegal content' should be defined broadly to cover information relating to illegal content, products, services and activities. In particular, that concept should be understood to refer to information, irrespective of its form, that under the applicable law is either itself illegal, such as illegal hate speech or terrorist content and unlawful discriminatory content, or that the applicable rules render illegal in view of the fact that it relates to illegal activities. Illustrative examples include the sharing of images depicting child sexual abuse, the unlawful non-consensual sharing of private images, online stalking, the sale of non-compliant or counterfeit products, the sale of products or the provision of services in infringement of consumer protection law, the non-authorised use of copyright protected material, the illegal offer of accommodation services or the illegal sale of live animals.

The definition of illegal content is thus very broad. It covers many fields of law and may also differ from one Member State to another. This raises the question how a hosting service provider is capable of interpreting all sections of criminal law, intellectual property rights, privacy and personal data regulation, compensation or tort law, consumer law and such special fields as tourist services and illegal sale of certain products. The concept of 'illegal content' is too difficult to interpret and requires high levels of expertise

and comprehensive knowledge of both EU and national legislation.[94] A broad obligation in terms of content removal can lead to a 'chilling effect' where lawful (but maybe harmful) material is removed for fear of civil lawsuits or criminal convictions.[95]

It remains to be seen how much the complaint-handling system will be used, because there are many publication channels, and if one service provider or platform administrator removes content, it is often possible to publish the same content elsewhere. It is also important to bear in mind that freedom of expression does not provide 'freedom of forum'.[96] In practice, however, a politician, for example, could suffer a serious setback if their Twitter or Facebook account with numerous followers were deleted, or if they could not be found via the Google search engine.

Content evaluation requires financial and intellectual resources which smaller content service providers in particular do not have. Measures taken should be necessary and proportionate.[97] Liability can also lead to some web site administrators not allowing readers to comment or only keeping comments open for a short time.[98]

Legal safeguards and the significance of communication technology

Content removal by platforms does not provide the same legal safeguards as court proceedings. Yet no one has an absolute right to express their views via a particular means of communication under the right to freedom of expression. That is why the administrators of a Facebook account, a blog, or the like can also decide on the content they publish or delete. However, the DSA will bring more rights and transparency regarding content removal policies and practices by hosting service providers. They may also fall under competition regulation if they are in a dominant market position.

When implementing regulatory solutions, it is important that moderation measures taken only to remove illegal content do not lead to liability and/or the service provider being considered as a publisher. An activity aimed at preventing the appearance of illegal content should always be more profitable than staying completely passive.[99] This is also manifest in the DSA's

[94] See, e.g, Lakivaliokunnan lausunto 9/2021 (n 9) and Perustuslakivaliokunnan lausunto 20/2021, 3 (statement of Finnish Constitutional Law Committee).

[95] Erixon (n 11), 1,4, 8,9.

[96] A Kuczerawy, 'Does Twitter trump Trump?' (Verfassungsblog, 29 January 2021) < www.verfassungsblog.de/twitter-trump-trump/> accessed 4 October 2022.

[97] See United Nations General Assembly, 'Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression' (2019) UN Doc A74/48050, 12–13.

[98] See D Rolph, 'Liability for third party comments on social media pages' (2021) 13 Journal of Media Law 122, 132-133.

[99] See also the Law Commission of Ontario, *Defamation Law in the Internet Age* (2020), 76–777, <www.lco-cdo.org/wp-content/uploads/2020/03/Defamation-Final-Report-Eng-FINAL-1.pdf> accessed 4 October 2022, where it is considered that 'the intermediary-as-publisher model encourages

Article 7, which stipulates that liability exemption should not be excluded solely because online service providers 'in good faith and in a diligent manner, carry out voluntary own-initiative investigations into, or take other measures aimed at detecting, identifying and removing, or disabling access to, illegal content, or take the necessary measures to comply with the requirements of Union law and national law in compliance with Union law.' This corresponds to the e-Commerce Directive, even though it is stated in a more complicated manner here. The same outcome would also be achieved with interpretation which is reasonable and takes into account the concept of 'caching', 'mere conduit' and 'hosting' and the aims of the regulation in question. To avoid difficulties in interpretating the regulation, it is a positive feature that it has been stated explicitly.

According to the DSA Article 52, the Commission could impose on very large online platforms fines not exceeding six per cent of their annual worldwide turnover in the preceding financial year. Such high fines might bring about a 'chilling effect' and too easily encourage content removal, especially when the service provider does not have a vested interest in publishing the content. For example, in GDPR Article 83 the maximum administrative fine is EUR 20 000 000, or four per cent of the total worldwide annual turnover of the preceding financial year, whichever is higher. In addition, the relationship between such liability regulation and criminal liability and the right not to be tried or punished twice *(ne bis in idem)* also calls for a more detailed assessment.[100]

Conclusion

This article demonstrates, based on the court practice of the CJEU, that the liability exemption of the e-Commerce Directive is not applicable to holders of web discussion forums. However, this as such does not of itself constitute liability, because a basis for liability must always be found in the national legal order.

A review of ECHR jurisprudence shows that the liability of website administrators for third party comments has arisen only exceptionally and, even then, mainly in cases of clearly unlawful hate speech and when the offending post has not been promptly removed. Additionally, conditions

intermediaries to avoid liability by avoiding knowledge of defamatory content. As a result, the model undermines corporate social responsibility by discouraging them from making proactive efforts to monitor and reduce defamatory content.'

[100]Lakivaliokunnan lausunto 9/2021 (n 9), 8. See also European Court of Human Rights: Guide on Article 4 of Protocol No. 7 to the European Convention on Human Rights; Right not to be tried or punished twice. Updated 30 April 2021, 9-10, and *Grande Stevens v. Italy v Hungary* App no 18640/10, 18647/10, 18663/10, 18668/10 and 18698/10 (ECtHR, 7 July 2014) where the Court found that heavy administrative fines imposed on the applicants were also "criminal" for the purposes of Arts. 6 and 4 of Protocol No. 7.

for criminal liability are more stringent than those for civil liability. Moreover, greater action is required from commercial business companies than from private individuals who run a website as a hobby.

The DSA does not directly apply to liability of website administrators, but rather applies to the intermediary services required to maintain a discussion site. A person who has been the target of hate speech can, under the DSA, ask the hosting service provider to remove illegal content when the website administrator does not respond or agree. The DSA offers the victim of illegal hate speech a faster and cheaper way to remove illegal offensive content than by filing a lawsuit. At the same time, however, the effects of the DSA on freedom of expression should be closely monitored. National courts, the CJEU and the ECtHR will also play an important role in finding the right balance between freedom of speech and the right to private life.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributor

Päivi Korpisaari is Professor of Communication Law at the Faculty of Law, University of Helsinki. Her main fields of research are media and communication law, personal data protection, tort law, constitutional law, human rights law, and criminal law concerning "freedom of expression offences". This work was supported by Helsingin Sanomain Säätiö (Helsingin Sanomat Foundation) under the project Viha ja julkisuus (Hate and publicity).