Social Network Users' Liability for Third-Party Comments and Potential Chilling Effects on Freedom of Expression

di Marco Bassini

Title: La responsabilità degli utenti dei social network per i commenti di terzi e i potenziali effetti di limitazione della libertà di espressione

Keywords: Content moderation; Freedom of expression; Art. 10 ECHR.

1. – Content moderation by private and profit-driven online platforms, such as social networks, has gained a prominent role in the current discussions on the most pressing challenges that the digital society poses to fundamental rights (E. Celeste et al., *The Content Governance Dilemma. Digital Constitutionalism, Social Media and the Search for a Global Standard*, London, 2023). There are debated issues in literature and case law, including whether users enjoy a "right to post" and whether freedom of expression applies "horizontally" to the contractual relationship between users and platforms performing content moderation (C. Bambrick, *Constitutionalizing the Private Sphere*, Cambridge, 2025; E. Celeste et al. (Eds), *Constitutionalising Social Media*, Oxford, 2022; M. Bassini, *Social networks as new public forums? Enforcing the rule of law in the digital environment*, in *IRIC*, 2022, 311). However, another key but comparatively underdeveloped perspective on the protection of freedom of expression "from" content moderation concerns the role of owners of public digital spaces, such as social network pages or websites, and their liability for moderating offensive third-party comments. While the ability of private platforms to perform content moderation on user-generated content is seen as "a matter of power" inherently connected with freedom of expression (A. Palumbo, *A Medley of Public and Private Power in DSA Content Moderation for Harmful but Legal Content: An Account of Transparency, Accountability and Redress Challenges*, in *JIPITEC*, 2024, 246; J.P. Quintais et. al, *Using Terms and Conditions to apply Fundamental Rights to Content Moderation*, in *Ger. Law J.*, 24(5), 2023, 881; A. Daly, *Private Power, Online Information Flows and EU Law: Mind The Gap*, Oxford, 2019), the status of webpage or account owners as moderators of third-party speech deserves careful investigation, since it has potentially chilling effects on this freedom. The judgment of the European Court of Human Rights in *Alexandru Pătraşcu v. Romania* seems to prove this point.

2. – The background of the case (T. McGonagle, *European Court of Human Rights: civil liability for publication and hosting of comments on Facebook violates freedom of expression*, in *IRIS*, 2025-3:1/11) does not significantly differ from that of many applications for an art. 10 of the European Convention of Human Rights-based review by the European Court of Human Rights. Alexandru Pătraşcu, a computer

engineer and passionate opera fan from Romania, used to actively comment on opera-related matters via his blog, "Sur l'opéra", and his Facebook page. He even published some articles in reputable national and international magazines. In 2016, a scandal broke out at the Bucharest National Opera (ONB) due to an internal conflict triggered by protests against foreign employees. For months, the applicant provided wide coverage of these developments. He opposed the protests and criticised their nationalist tone, frequently posting about the situation, including remarks on two active members of the protests who were appointed managerial roles by the Minister of Culture. As a result, Mr. Pătraşcu's publications attracted several third-party comments that were posted on his blog and his Facebook wall. A few months later, the two employees who participated in the protests brought a civil lawsuit against the applicant for defamation. They requested the removal of defamatory comments from his platforms and sought compensation for the harm to their reputation and dignity. The Court of Bucharest held the applicant liable not just for his own posts but also for failing to moderate defamatory comments made by third parties on his blog and his Facebook page. The court equated Pătraşcu's role to that of a TV moderator, holding him responsible for the failure to remove many comments that were harmful to the plaintiffs' reputation and honour. The applicant was, accordingly, ordered to pay over 8,000 euros for non-material damages, and he was required to prevent the publication of offensive comments by third parties on his blog and Facebook account. In its judgment, the court of first instance did not differentiate the content posted by the applicant and the content posted by third parties on the applicant's blog and Facebook wall. The Court of Appeal partially overturned the first instance decision. It narrowed the scope of the content for which the applicant was found liable down to 22 comments, including four posted by Pătraşcu himself, which were found to be offensive. In the Court of Appeal's reasoning, although the owners of social network pages are not required to conduct any prior monitoring of the content posted by third parties, they are obliged to act after being notified that the offensive content has been posted. In the case at hand, the Court of Appeal found that the applicant had incurred liability because of this inaction, having been informed of the offensive nature of some comments but tolerating them. The appellate court, however, overturned the order for the applicant to refrain from publishing offensive third-party comments in the future, as such a negative obligation could not be imposed by a court of law. Ultimately, the Court of Appeal significantly reduced the amount of compensation owed for damages – the amount was dropped to approximately 2,500 euros. The appellate court's decision was later upheld by the High Court of Cassation and Justice, before which the applicant emphasised that a finding of liability would conflict with art. 10 ECHR. The High Court did not engage in an in-depth review of the merits of the case. However, the High Court did point out that, on one hand, the duty to remove third-party content after notification was consistent with the European Court of Human Rights' case law, and more specifically with the landmark precedent in the *Delfi v. Estonia* case (ECtHR, *Delfi v. Estonia*, app. no. 64569/09, 16 June 2015; for some comments, see N. Cox, *Delfi AS v Estonia: The Liability of Secondary Internet Publishers for Violation of Reputational Rights under the European Convention on Human Rights*, in *Mod. L. Rev.*, 77(4), 2014, 619; D. Voorhoof, *Delfi AS v. Estonia: Grand Chamber confirms liability of online news portal for offensive comments posted by its readers*, in *Strasbourg observers*, 18 June 2015, available at strasbourgobservers.com/2015/06/18/delfi-as-v-estonia-grand-chamber-confirms-liability-of-online-news-portal-for-offensive-comments-posted-by-its-readers). On the other hand, the High Court also pointed out that Mr. Pătraşcu's role as a cultural commentator made him a content provider, and not a merely passive user. Therefore, the applicant brought a complaint based on the alleged violation of his right to freedom of expression under art. 10 ECHR.

ISSN: 2037-6677

3. – In its 7 January 2025 decision, the European Court of Human Rights found a violation of art. 10 ECHR on two distinct grounds: (1) the applicant's civil liability for his own statements posted on Facebook and (2) the finding of civil liability for failing to remove comments posted by third parties on his page. The judgment of the European Court is especially interesting and innovative due to the second line of reasoning. As far as the applicant's liability for his own statements was concerned, the European Court engaged in its traditional art. 10-based review of the domestic judgments. The Court found that, however provocative and offensive, the expressions of the applicant contributed to a debate of public interest regarding the governance of a major cultural institution. Equally, the Court emphasised the domestic courts' failure to distinguish between value judgments and factual statements and their failure to demonstrate that the restriction was justified by a pressing social need. Accordingly, the European Court declared that the interference posed by the Romanian authorities on the applicant's right to freedom of expression was disproportionate. The most interesting part of the reasoning of the European Court, however, concerns the second ground, *i.e.*, the finding of civil liability for the comments posted by third parties. While the scrutiny carried out by the Court regarding the first line of reasoning was rooted on the third pillar – *i.e.*, the proportionality of the restriction – the rationale of the Court on this second point has moved from the requirement of legal foreseeability of the restriction. As a matter of fact, the European Court found that the general provisions on tort liability of the Romanian Civil Code, on which the domestic courts had relied (namely, art. 1349 and art. 1357), could not amount to a proper legal basis for the challenged restriction. In the view of the Court and as a general expression of a *neminem laedere* duty, art. 1349 and art. 1357 could not *per se* meet the foreseeability requirement: These provisions do not define any specific duty for individual users of social network platforms to manage their own pages and monitor interactions of third parties. The European Court noted that there were no established precedents in national case law nor legal provisions that could have reasonably made the applicant aware of the existence of such a duty, and especially the obligation to remove third-party comments in the presence of a lawsuit. Moreover, the European Court found that Facebook's own terms of service did not impose any duty to delete controversial content. As a result, the Court found that the restriction did not meet the requirement of foreseeability, and thus the principle of legal certainty. But the judgment also adds key clarification regarding the reasoning of the Romanian High Court, comparing the specific case to the *Delfi v. Estonia* landmark judgment. In this case, the European Court had found the imposition of a low fine for failure to remove third-party comments posted on the webpages hosted by the news portal to be compatible with art. 10 ECHR. In the judgment under comment the European Court made it clear that the two scenarios are quite different. As a matter of fact, Delfi was a profit-driven news portal that exercised editorial control over the content it published. In the case at hand, however, the applicant was merely a private user of a social network platform – a platform which he had no responsibility to personally moderate. In addition, he could have limited control on the comments posted by third parties, and he might not solicit or endorse them. In the end, the European Court found that no prior request to remove the comments in question had been made before the litigation. Similarly, the comments had not been neglected but just preserved for evidentiary purposes. Under these circumstances, imposing liability on the applicant in his role as user of a social network amounted to an unforeseeable restriction not prescribed by law. In a nutshell, while the judgment of the European Court held the finding of the applicant's liability for his own statements a disproportionate interference with his freedom of expression, it considered a violation of art. 10 ECHR the finding of civil liability for third-party comments due to the lack of foreseeability of the restriction.

4. – The added value of the European Court's judgment can be understood in light of the Courts' prior cases concerning liability for third-party comments posted online (R. Spano, *Intermediary Liability for Online User Comments under the European Convention on Human Rights*, in *Hum. Rights Law Rev.*, 17(4), 2017, 665). Although content moderation is generally conceptualised as an activity that is primarily governed by contractual terms of service and relevant in the context of private platforms, the far-reaching implications on fundamental rights, and especially on freedom of expression, have been well-captured in the case law of the European Court of Human Rights. The Court's precedents have mainly focused on the role of online intermediaries, with a focus on assessing the compatibility with art. 10 ECHR of findings of civil liability by domestic authorities depending on the performance of content moderation. The case law of the Court has clearly pointed out that, by imposing fines or damage compensation for the failure to properly carry out content moderation, state authorities may interfere with freedom of expression. However, this has not prevented the Court from engaging in its three-pronged review of the lawfulness, legitimacy, and necessity of each interference. The point was made first in the leading *Delfi v. Estonia* judgment. In this case, the Grand Chamber found that the imposition of moderate damages due to the failure of an online news portal to promptly remove third-party comments amounting to hate speech constituted a justified and proportionate interference, which therefore did not violate art. 10 ECHR. As a result, the judgment of the Grand Chamber essentially framed the question as a balance of interests, but interestingly it accepted that the platform could be liable for failing to moderate content even before the latter was notified by the victims (M. Husovec et al., *Grand confusion after Sanchez v. France: Seven reasons for concern about Strasbourg jurisprudence on intermediaries*, in *Maastricht J. Eur. & Comp. L.*, 31(3), 2024, 385, 388). This finding raised significant criticism for various reasons. First, in most of the jurisdictions of the Contracting States and in all the European Union Member States, online intermediaries do not monitor nor filter third-party content, even if such intermediaries separately exercise some control over the content they publish under their editorial responsibility, as in the case of Delfi. Platforms like social networks – which are governed in the EU by the Digital Services Act (DSA) (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, and prior to that by the E-Commerce Directive, Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, especially electronic commerce, in the Internal Market) – are primarily framed as service providers acting as intermediaries and face liability for third-party content only when they fail to comply with a notice-and-action mechanism. This mechanism plays a key role, preventing intermediaries from incurring any liability for third-party content unless (and until) they have actual knowledge or awareness about the allegedly unlawful item of information concerned (for an overview, E. Apa-O. Pollicino, *Modelling the Liability of Internet Service Providers. Google vs. Vivi Down. A Constitutional Perspective*, Milan, 2013; M. Bassini, *Fundamental rights and private enforcement in the digital age*, in *Eur. Law J.*, 25(2), 2017, 182). Through the notice-and-action mechanism (which replaced the previous "notice-and-take down" scheme), modelled by its art. 16, the DSA determines the circumstances under which intermediaries are presumed to have such knowledge or awareness, marking the moment from which they can incur liability for third-party content if they fail to take illegal content down. In light of the EU legal framework, it becomes clear that the judgment in *Delfi v. Estonia* was not entirely consistent with the assumptions behind the E-Commerce Directive and, now, the Digital Services Act. As noted by some scholars (M. Husovec et al., *Grand confusion after Sanchez v.*

France, cit., 388), the judgment seemed to imply that States «would be free to impose monitoring obligations on platforms concerning third-party content». However, the Court also showed that the case concerned the duties and responsibilities of internet news portals providing, for economic purposes, a platform for user-generated comments on previously published content. Moreover, it did not extend to other *fora* on the internet, such as social media platforms, where the provider does not offer any content, and the content provider may be a private user (§§ 115-116). A second, intertwined critique centred on the risk of collateral censorship (J.M. Balkin, *Old-School/New-School Speech Regulation*, 127 *Harvard Law Review* 2296 (2014)) was voiced by judges Sajó and Tsotsoria in their dissenting opinion. Sajó and Tsotsoria highlighted how the Court «approved a liability system that imposes a requirement of constructive knowledge on active Internet intermediaries» (§ 1), likely to bring undesirable consequences from a freedom of expression perspective. As the two judges put it: «For the sake of preventing defamation of all kinds, and perhaps all "illegal" activities, all comments will have to be monitored from the moment they are posted. As a consequence, active intermediaries and blog operators will have considerable incentives to discontinue offering a comments feature, and the fear of liability may lead to additional self-censorship by operators. This is an invitation to self-censorship at its worst» (§ 1). Only one year later, the European Court partially revisited its approach in the *MTE v. Hungary* case (ECtHR, *Magyar Tartalomszolgáltatók Egyesülete and Index.hu Zrt v. Hungary*, app. no. 22947/13, 2 February 2016; for some comments, see C. Angelopoulos, *MTE v Hungary : A new ECtHR judgment on intermediary liability and freedom of expression*, in *JIPLP*, 11(8), 2016, 582; P.J. Ombelet-A. Kuczerawy, *Delfi revisited: the MTE-Index.hu v. Hungary case*, in *LSE Media Blog*, 19 February 2016, available at blogs.lse.ac.uk/medialse/2016/02/19/delfi-revisited-the-mte-index-hu-v-hungary-case), where it further elaborated on the intersection between norms governing liability of online intermediaries and possible chilling effects for freedom of expression. On this occasion, the European Court tried to limit the effects of its previous stance by emphasising that the notice-and-take down mechanism could serve as a proper solution for balancing the various interests at stake (*e.g.*, the protection of freedom of expression and that of individuals' reputation), as long as there is no clearly unlawful content (§ 91). As a matter of fact, the Court found that the order to pay damages for third-party defamatory comments that were imposed to an online news portal by the Hungarian authorities, reflecting a strict liability standard, would undermine the right to impart information on the internet, thus violating art. 10 ECHR. In the attempt to justify its *distinguo* with the decision taken in *Delfi v. Estonia*, the Court also underscored that, in the latter case and unlike in *MTE v. Hungary*, the anonymous comments posted by the portal readers amounted to hate speech. The Court could therefore fine-tune the set of criteria to review the compatibility with art. 10 ECHR of the finding of liability for third-party comments: (i) the context and content of the comments, (ii) the measures applied by the applicant company in order to prevent or remove defamatory comments, (iii) the liability of the actual authors of the comments as an alternative to the intermediary's liability, and (iv) the consequences of the domestic proceedings for the injured parties and for the applicants. The Court has consistently applied these criteria in subsequent cases regarding intermediaries' liability for third-party comments (ECtHR, *Pihl v. Sweden*, app. no. 74742/14, 9 March 2017; ECtHR, *Tamiz v. UK*, app. no. 3877/14, 19 September 2017; ECtHR, *Høiness v. Norway*, app. no. 43624/14, 19 March 2019).

5. – A remarkable development in the case law of the European Court of Human Rights came in 2023 with the *Sanchez v. France* judgment (ECtHR, *Sanchez v. France*, app. no. 45581/15, 15 May 2023; for some comments, see J. van de Kerkhof, *Sanchez v France: The Expansion of Intermediary Liability in the Context of Online Hate Speech*, in *Strasbourg observer*, 17 July 2023, available at strasbourgobservers.com/2023/07/17/sanchez-v-france-the-expansion-of-intermediary-liability-in-the-context-of-online-hate-speech; D. Voorhoof, *Sanchez v. France: Grand Chamber confirms decision that criminal conviction for hate speech does not violate Article 10 ECHR*, in *Inforrm's Blog, The International Forum for Responsible Media Blog*, 2023), which stands out as a key term of reference to understand the impact of the case concerned. The Grand Chamber dismissed the application of a French politician who had been convicted for failing to remove racist and xenophobic comments posted by third parties on his public Facebook wall. From the perspective of the Court, the criminal conviction imposed by the French courts constituted a lawful, legitimate, and proportionate restriction on the applicant's freedom of expression. The judgment attached significance to some circumstances that were taken into account by the French authorities in the domestic proceedings, including the fact that the applicant had not taken action to moderate offensive comments for weeks following their publication, thus failing to prevent their dissemination. The French courts highlighted that the applicant had deliberately chosen not to restrict users' access to his public Facebook wall, and that this decision gave rise to specific obligations. They reasoned that the applicant had to be considered liable as the producer of the website where the third-party comments were posted. Although the liability for such comments should only arise when the producer had knowledge of their content before they were posted, which was not the case of the applicant, the French courts stressed that he was a public figure and the fact that he had decided to make his Facebook wall public made him responsible for the comments posted. The applicant could not claim that he was not aware of the existence of clearly unlawful content. As noted (M. Husovec et al., *Grand confusion after Sanchez v. France*, cit., 395), in the *Delfi v. Estonia* and *MTE v. Hungary* cases the European Court seems to have valued a specific connection between the nature of the entity or individual expected to engage in content moderation and the applicable liability standard: As Husovec and others have claimed (M. Husovec et al., *Grand confusion after Sanchez v. France*, cit., 395), the Court developed a "professional criteria" to ascertain whether the applicant had a commercial structure and could presume a certain level of knowledge of the applicable law and the risks inherent in the relevant business. Without prejudice to the different type of content involved (hate speech in the former case, statements harmful to a company's business reputation in the latter), it is perhaps not coincidence, as the authors note, that the application of Delfi (a profit-driven company) was dismissed while that of MTE (a non-profit entity) was successful. After all, there was a difference in magnitude for the potential chilling effects that were derived from a finding of liability. What the judgment of the European Court in *Sanchez v. France* seems to do is extend the consideration as "facilitator" to a public figure, on the assumption that a politician should be well-aware of the implications of entertaining a public social network wall in the context of an electoral process and the consequences of its use as a means of political propaganda.

6. – The judgment in *Pătraşcu v. Romania* gave the European Court of Human Rights a new opportunity to elaborate on the compatibility with art. 10 ECHR of the liability of the owner of a public Facebook wall for third-party unlawful comments. Against the background of the *Delfi v. Estonia*, *MTE v. Hungary*, and *Sanchez v. France* cases, the Court could address various points to clarify the

circumstances under which a finding of liability would be compatible or not with art. 10 ECHR (K. Lemmens, *"That's what she said!" – Alexandru Pătraşcu v. Romania*, in *Strasbourg observers*, 11 April 2025, available at strasbourgobservers.com/2025/04/11/thats-what-she-said-alexandru-patrascu-v-romania). The European Court could, for example, provide clarification regarding the status of an individual that, albeit not amounting – at least *prima facie* – to a public figure, had a quite remarkable audience and was an influential national commentator on opera. On one hand, the applicant did not have any commercial purpose or structure, acting as merely a passionate individual interested in opera, then in a non-professional capacity and for no profit (like the applicant in the *MTE v. Hungary* case). On the other hand, however, it could be questioned whether, under these circumstances, he could nonetheless qualify as a facilitator (like the applicant in the *Sanchez v. France* case). This preliminary but essential point could pave the way for investigating the role of a private social network user in the context of moderating third-party comments. The Court could have applied the criteria developed in *Delfi v. Estonia* and further refined in *MTE v. Hungary* to the specific scenario where offensive comments were posted by third parties, or the Court could have drawn some connections to *Sanchez v. France*. Furthermore, the case did not concern the moderation of comments that themselves constituted hate speech, unlike in *Delfi v. Estonia* and *Sanchez v. France*. Because of the peculiar scenario, the case did not perfectly match any of these three prior cases, yet it featured important similarities with all of them. Another key point that the judgment could have addressed is the liability of social media account holders regarding comments that are posted by others, and how disputes surrounding this matter can arise. The Court had already acknowledged a lack of consensus on this point among the Contracting States in *Sanchez v. France* (§ 79), depending in most of the jurisdiction on the lack of a specific legislative choice. In the end, the judgment of the European Court has only focussed on the requirement of quality of the law providing for the restriction on freedom of expression. As noted above, in the domestic proceedings, the Romanian courts had enforced two general provisions on tort liability, namely, art. 1349 and art. 1357 of the Civil Code. The European Court noted that these provisions do not imply any obligation on the part of the applicant, as the owner of a Facebook page, to monitor content posted on this page by third parties (§ 128). Moreover, from the perspective of the Court, they do not provide any details as to the circumstances under which the owner of such a page might be required to carry out such monitoring, nor do they provide details about the measures that should be taken as a result and the conditions that may determine their fault. The European Court could not help but noting that the language of these provisions is extremely general and makes no reference to the specific social network platforms or to the context of content moderation. The Court did not immediately rule out these provisions, acknowledging that sometimes law must employ a general language. However, it tested them in their practical application, finding that indeed each of the domestic courts had ruled on the applicant's liability for third-party comments on seemingly different grounds. In more detail, the European Court pointed out that the court of first instance had presumed the applicant's liability on the assumption that his role was comparable to a moderator. On the contrary, the Court of Appeal had found the applicant liable for violating the terms of service of the social network, which, from its perspective, imposed on the user the same censorship obligation as that applicable on the platform. Finally, the High Romanian Court, applying the *Delfi v. Estonia* standard to the case, had suggested that the applicant was to be equated with a content provider, as this would have required the applicant to preventively engage in content moderation for third-party comments regardless of prior notice by the concerned party or by the social network. Based on the overall analysis of these judgments, the European Court

found that, given the different paths followed in the domestic courts' reasoning (which went far beyond the literal meaning of the enforced provisions), a finding of liability of the applicant could all but constitute a case law creation. That said, the European Court noted that this case law could not, at the time, provide any precise and coherent legal basis for holding the applicant liable for the third-party comments. From the perspective of the Court, if the legal basis for the restriction on the applicant's freedom of expression were framed this way, there would be no sufficient safeguards to protect individuals from the possible interferences by state authorities. The claimed legal basis, as a matter of fact, did not define the scope and modalities of the interference in the exercise of freedom of expression by an individual, and what that meant specifically through the opening of a social network account with sufficient clarity to allow such a user to enjoy a degree of protection of his freedom in a democratic society. Given the remarkable premises of the case and the landmark precedents of the Court, the judgment has only partially met the high expectations raised by the background of the case. The European Court has failed to provide further clarification and draw more precise boundaries between different possible scenarios that may trigger equally different conclusions about the liability of the owner of a social network page. Ultimately, the Court did not exclude that an interpretation of the role of the applicant like those adopted in the domestic courts' reasoning may be compatible with art. 10 ECHR. It did only exclude that such an interpretation could rely on the general provisions on tort liability enforced by the Romanian courts.

Marco Bassini
Tilburg Law School
Tilburg University
m.bassini@tilburguniversity.edu