PERCEPTIONS

Policy Brief

July, 2022

LIABILITY OF INTERNET SERVICE PROVIDERS FOR HATE SPEECH IN MIGRATION CONTEXTS

Jesus Ivan Mora Gonzalez

University of Granada

Executive Summary

The purpose of this policy brief is to analyse the liability of internet service providers in cases of hate speech directed towards the migrant population, with the aim of protecting all parties on the internet. In the European Union, liability of internet service providers is regulated by the E-Commerce Directive and covers several services, such as transmission, catching, or data hosting, which are key for platforms such as Facebook, Twitter, Youtube, Amazon, or Ebay. Internet users rely on these services, but such services can also be used to commit crimes. This situation requires a deep study on liability to develop clear normative criteria for cases where platforms are used by third parties. However, liability must be established in a balanced way, taking into account the fundamental rights assessment of the different parties involved.

Introduction

The prevalence of internet service providers that primarily offer a data hosting service through which third parties can express their opinion, such as Facebook, Whatsapp or Twitter (among others), has created a new dimension for hate speech. Such platforms have become fundamental agents for the promotion of freedom of speech, but also for the protection of social groups most targeted by hate speech or the dissemination of fake news (in this case the migrant population). It is therefore necessary to identify hate speech on the main social platforms and to decide upon the liability of both perpetrators of hate speech and of the secondary parties that allow it to happen: internet service providers. The importance of this point is highlighted in the PERCEPTIONS baseline report (D2.6, Lore Van Praag & Rut Van Caudenberg, 2020), which emphasises the need to delimit the liability of internet service providers, and reflects upon how host society groups can become essential players in the protection of fundamental rights, especially in the prevention of threats to migrants like discrimination or hate speech.

One illustrative example of hate speech on the abovementioned platforms is the case of Twitter and news regarding the arrival of the Aquarius boat to the Spanish coast. Between the 8th of June of 2018 and the 17th of June 2018, there were around 24.000 messages on Twitter related to the Aquarius' arrival, of which 27% could potentially be regarded as hate messages directed at the migrant population (and able to influence the general population's perceptions of this group). Such a situation is problematic because it not only requires a response to the third parties who commit the crime as perpetrators (primary liability), but also to interest service providers, such as Twitter, for allowing this hate speech to happen.

Key Issues:

- *Hate Speech*
- *Internet Service Providers*
- *Secondary Liability*

Analysis

In the European Union, the key regulation of the liability of internet service provider in cases of hate speech is the Directive 2018/1808, concerning the provision of audio-visual media services. This directive makes a very brief mention of taking appropriate measures regarding hate speech online, but taking into account the normative model developed by the Electronic Commerce Directive (2000), which is known as "notice and takedown". According to this model, internet service providers, such as Twitter, will not be liable (safe harbour) as long as, after notification of content infringement, they remove the content. However, this model ignores the problem of balancing an efficient process with a due process rights for all parties involved. For the proper assessment of all interests, it is important to develop a legal model that recognises this balanced approach, which should take into account two basic variables. Firstly, the internet service provider's capability to identify and assess the infringement committed by a third party as perpetrator (primary liability) and, secondly, the defence's right, as a due process right, prior to the

Key Findings:

- *European Regulation*
- *Notice and Takedown*
- *Contributory Liability*

blocking of access to the content. This would imply the legal recognition of some scope for no collaboration by the internet service providers, without them losing the status of neutral actor to protect freedom of speech. At the same time, however, it would make the service providers responsible for this infringement of content (hate speech) if, after hearing all parties, they do not remove said content. That is, if the third party is not able to justify this speech, the internet service providers must act to protect the interest of groups subject to this hate speech.

To accommodate these interests on the internet, the concepts of omissive participation as contributory liability (secondary liability) and the guarantor position appear as normative criteria that could facilitate progress towards European harmonization of internet service liability, moving towards a procedure able to deal with the problem of hate speech against migrant population, whilst respecting the fundamental rights of all parties. A key point, in this sense, would be to consider omissive behaviour from internet services providers as active conduct contributing to the infringement committed by a third party. That is, the internet service provider would be seen to contribute to the unlawful act committed by a third party, which is not blocked when there are sufficient proofs derived from the hearings of all parties (expressing a mere potential and negative dominion of the parties). This is important because understanding the internet service providers as supervisors of the main risk committed by third parties would imply equating primary and secondary liability, legitimising preventive models for the removal of allegedly infringing content by a general supervision duty, which is forbidden by Article 15 E-Commerce Directive.

Therefore, contributory liability provides an interpretative approach able to distinguish between third parties who commit the hate speech (primary liability) and secondary parties (internet service providers) who would be obliged to control a risk that is only indirectly realised on the third-party behaviour. It would reflect potential control, rather than direct control, with regard to the third-party speech that would require, at the same time, internet service providers having specific knowledge (mens rea) of the infringement committed by the third party. Such knowledge would not be required if the internet service provider were considered the primary party. In such a situation, the primary party would be subject to direct liability, in which due diligence would make it possible to derive internet service providers' mens rea from general knowledge of possible infringements that take place on the platform organised by the intermediaries.

Recommendations

A harmonised model in the European Union regarding the liability of internet service providers for hate speech (which has the potential to affect both the migrant population and the general population's perception of this group) should contain at least the following requirements:

A. Formal requirement to justify specific knowledge regarding a specific instance of hate speech: internet service providers' capability as a mere mediator

The recognition of internet service providers as mere mediators, and therefore, exempted from secondary liability, should be justified by their inability to implement a specific duty of protection consisting of preventing the continuation of the infringing activity by third parties. From this point of view, the affected parties or victims should notify internet service providers of the hate speech committed and the reasons for which they think such an offence exists. The concept of affected parties should be extended not only to specific victims targeted by this hate speech, but also to NGOs or other organisations with specific interests in the protection of these groups, and state bodies responsible for the prosecution of crimes or infringements. If justification of the offence is not explained to the internet service provider, the provider's role should be reduced to the condition of mere mediator, informing the affected parties of the possible infringement, with state bodies responsible for the prosecution as the best positioned actors to investigate the case. At the same time, the refusal by the internet service provider to remove or block the content should not imply liability, as it would be impossible to assess the case correctly due to the different fundamental rights at stake.

B. Specific knowledge of the infringement

Once internet service providers' capability to examine the infringement (in this case, hate speech) have been accepted, they should be obliged to implement a specific duty as guarantor, consisting of notifying the alleged offence committed by a third party and identifying the counter-notification right to justify content uploaded to the internet. This counter-notification right should be mandatory in order to guarantee a defence right (due process) prior to the removal or blocking of any content and to determine the specific knowledge of the infringement as a basis upon which to justify internet service providers' liability. If there is a counter-notification by the alleged infringer, and if, from this counter-notification, potentially legitimate content can be justified, the internet service provider should contact the affected parties again to inform them that the content has not been removed or blocked, and that their case should be conducted through a judicial process, losing the condition of neutral actor (safe harbour).

Key recommendations:

- *Specific Knowledge*
- *Internet Service Providers' Capability*
- *Due Process*
- *Counter-notification*
- *Safe Harbour*
- *Guarantor Position*

C. Breach of the duty to prevent the continuation of the unlawful act: liability scope of internet service providers

The liability of internet service providers should be delimited through an omissive behaviour generated by specific knowledge derived from counter-notification, which makes it possible to establish the equivalence between the omissive behaviour of internet service providers and the active conduct of contribution to the dissemination of infringing content by a third party. This will be possible as long as the counter-notification shows the acceptance of the infringement by the third party, there is no coherent justification for the content posted on the internet, or the alleged infringer's response is non-existent. As a result, internet service providers will be only responsible as a contributor (secondary liability) if they decide not to collaborate when counter-notification clearly shows the existence of an infringement that internet service providers have the capacity to remove or block.

References

Angelopoulos, C & Smet, J (2016) Notice and Fair Balance: how to reach a compromise between fundamental rights in European Intermediary Liability, Journal of Media Law, 8, 1-26.

Arcila-Calderon, C et all (2021) Refugees Welcome? Online Hate Speech and Sentiments in Twitter in Spain during the reception of the Boat Aquarius. Sustainability, 13, 1-16.

Bruner, L (2016) The Liability of an Online Intermediary/Third Party Content: the Watchdog becomes the Monitor: Intermediary Liability After Delfi. Human Rights Law Review, 16, 163-174

García-Carmona, M, García Quero, F.J., Guardiola, J., Moya Fernández, P., Ollero Perán, J., Edwards, J., & Withworth, B (2021). Migration to the EU: a survey of first-line practitioners' perceptions during the COVID-19 pandemic.

Helberger, N (2017) Governing Online Platforms: From Contest to Cooperative Responsibility. The Information Society, 34, 1-14

Keats Citron, D (2008). Technological Due Process. Washington University Review, 85, 1249-1314.

Longke, T (2019) On an Internet Service Provider's Content Management Obligation and Criminal Liability. Journal of Eastern-European Criminal Law, 1, 145-158.

Montagnani, M & Yordanova, A (2018) Safe Harbors in Deep Waters: A New Emerging Liability Regime for Internet Intermediaries in the Digital Single Market. International Journal and Law and Information Technology, 26, 294-310.

Prechal, S (2020) Horizontal Direct Effect of the Charter of Fundamental Rights of the EU. Revista de Derecho Comunitario Europeo, 66, 408-426.

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'). Available at: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex%3A32000L0031

Directive (EU) 2018/1808 of the European Parliament and of the Council of 14 November 2018 amending Directive 2010/13/EU on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive) in view of changing market realities. Available at: https://eur-lex.europa.eu/eli/dir/2018/1808/oj

○ Websites

www.perceptions.eu

www.project.perceptions.eu

○ Deliverables

Bayerl, S., Pannocchia, D., & Hough, K. (2019) Deliverable D2.2 Secondary Analysis of studies, projects and narratives. PERCEPTIONS H2020 Project No 833870

○ Contact

Office@perceptions.eu

jesusivanmora@ugr.es

Acknowledgement: This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 833870.