CENTRAL EUROPEAN UNIVERSITY

ONLINE HATE SPEECH: AGAINST STRICT INTERNET INTERMEDIARIES' LIABILITY AS A SOLUTION FOR ENSURING VICTIM'S PROTECTION

By Milica Nešić

LL.M Thesis
PROFESSOR: Andras Sajó
Central European University
1051 Budapest, Nador utca 9,
Hungary

TABLE OF CONTENTS

EXECUTIVE SUMMARY

As a response to online hate speech, countries started introducing strict liability regime for Internet Intermediaries who host third-party content. In the recent years, this trend of obliging Internet Intermediaries to monitor all the content they host in order to expeditiously delete illegal one, has started spreading especially among European countries.

The thesis analyzes the strict liability regime imposed to Internet Intermediaries for third-party content in the cases of online hate speech, by presenting solutions adopted recently by ECtHR and Germany. At the same time, it makes parallel to the Indian solution, where broad immunity regime was introduced, making Internet Intermediaries liable only in case they fail to obey a court order.

Seeking to provide an answer to the question whether imposing strict liability to Internet Intermediaries is a proper way of combating online hate speech, the thesis' major finding concludes that Internet Intermediaries, as private actors, should not be imposed with such type of liability, especially not in the cases of online hate speech. This regime undermines freedom of expression and introduces high level of censorship, since Internet Intermediaries, as private actors lacking legal knowledge and afraid of being punished, often err and delete content out of caution, especially when deciding about alleged hate speech content. At the same time, strict liability regime represents an excessive burden for Internet Intermediaries, since it requires additional resources and enormous personnel in order to pre-screen entire third-party content. Additionally, it endangers third-parties due process rights and privacy rights.

By explaining all the negative implications of strict liability regime through comparative analysis of different legal solutions, the thesis in a systematic manner calls upon the states to refrain themselves from imposing this type of liability to Internet Intermediaries.

LIST OF ABBREVIATIONS

ACHR – American Convention on Human Rights

Act to Improve Enforcement of the Law in Social Networks – Network Enforcement Act

AfCHR – African Charter on Human and Peoples' Rights

CJEU – European Union Court of Justice

Committee of Ministers – Committee of Minsters of the Council of Europe

Genocide Convention – Convention on the Prevention and Punishment of the Crime of Genocide

ICCPR – International Covenant on Civil and Political Rights

ICERD – International Convention on Elimination of Racial Discrimination

Information Technology Rules – Information Guidelines

IT Act – Information Technology Act

The Convention, the ECHR – The European Convention on Human Rights

The Court, the ECtHR – The European Court of Human Rights

The E-Commerce Directive – The E-Commerce Directive 2000/31/EC

UNESCO – United Nations Educational, Scientific and Cultural Organization

UNGA – United Nations General Assembly

UDHR – Universal Declaration of Human Rights

UNHRC – United Nations Human Rights Council

INTRODUCTION

During the past few years, the usage of the Internet has increased widely. At the moment, the number of Internet users has reached striking 3,424.971.237, which represents 46.1% of the whole world population.[1] Spending time in online space has become a part of people's everyday life. Internet has been extremely beneficial to the mankind. It offers a possibility of accessing different information making it easier for people to enrich their knowledge bank. Most importantly, people have been offered with an opportunity to share their views freely and exchange their ideas and opinions, contributing to the "market of ideas". At the same time, Internet has increased wide-open debates on different public issues, which undoubtedly represents a crucial part of every free and open society today.

Although Internet has been beneficial for people in many ways, at the same time it has been misused for different purposes. Internet has become, as well, a forum for incitement to violence and hatred, since it has provided the possibility for the perpetrators, while performing the crime, to hide their identity easily and avoid facing punishment. Many authors started emphasizing the severity of online hate speech in comparison to the offline one by pointing out that hateful comments can go viral, beyond any possible border and by unimaginable speed.[2] They were also adding that the perpetrator was often anonymous[3], thus, preventing victims from being remedied.[4]

As Internet is being used for these purposes as well, online hate speech has become a problem that, according to many international actors, required action. International organizations and human

[1] This number was almost seven times smaller in year 2000; for further statistical data see <http://www.internetlivestats.com/internet-users/> accessed 2 December 2017
[2] Dragos Cucereanu, *Aspects of Regulating Freedom of Expression on Internet* (1st edn, Intersentia, 2008), 193
[3] ibid (n 2), 183
[4] Tarlach McGonagle, 'The Council of Europe against online hate speech: Conundrums and challenges' p 29 – 30 <https://rm.coe.int/16800c170f> accessed 13 October 2017

rights bodies, as well as national legislators, have been trying to adopt effective legal documents and provide adequate measures in order to properly address the question of online hate speech. However, one of the biggest problem they encounter is finding a proper solution for detecting and removing illegal hateful content.[5] As a result, a trend to include Internet Intermediaries into the whole protection picture has been developed, since according to many – as the ones who host third-party content which may represent hate speech and as the ones who have enough financial and technical means to block or remove hateful content[6] – Internet Intermediaries are seen as crucial players in the hate speech suppression.

However, there is still no harmonized approach about the role of the Internet Intermediaries in countering online hate speech and different jurisdictions address this issue differently.[7] Usually, Internet Intermediaries, are being seen just as "messengers" or "hosts", and since they do not modify the allegedly unlawful content they host, they cannot be held liable.[8] However, for some international bodies and national legislators, Internet Intermediaries could be held liable for unlawful third-party content, and level of their liability may differ.

In order to combat online hate speech, Governments started putting pressure on Internet Intermediaries calling for their self-regulation. In that manner, in 2016, big Internet Intermediaries such as *inter alia* Facebook, Twitter and YouTube formulated a Code of Conduct on countering illegal hate, committing themselves to, beside other commitments, remove online hate speech in

[5] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, 'Tackling Illegal Content Online: Towards an enhanced responsibility of online platforms' COM (2017) 555, p 2

[6] Article 19, 'Internet Intermediaries: Dilemma of Liability' (2013) p 8 <https://www.article19.org/wp-content/uploads/2018/02/Intermediaries_ENGLISH.pdf> accessed 7 February 2018

[7] Francesco Buffa, *Freedom of expression in the Internet society* (Key, 2016), 34

[8] Article 19 (n 6), p 2

less than 24 hours.[9] Although not imposing liability to Internet Intermediaries, but relying on their good faith leaving them the opportunity to set up their own self-regulatory mechanisms, this position is nothing more than the reflection of the idea where Intermediaries, as private entities, are left to decide which content is illegal and which is not.

Unfortunately, instead of pushing for self-regulation, for some actors this problem was addressed by introducing solutions which were detrimental to human rights. Namely, faced with society's fear from being victimized which spread even more after the birth of the Internet, States found themselves trapped and rushed to offer quick solutions which would provide security to their people. However, such blast decisions, consisting from introduction of strict liability regimes for Internet Intermediaries in the cases of online hate speech, did not take into the account the human rights that everyone were entitled to and that should not be taken for granted. States started imposing strict liability regimes by introducing laws which require from Internet Intermediaries to monitor the content they host and delete if expeditiously enough if they do not want to face huge fines. Exactly this approach has begun to spread among different jurisdictions and legislators have started to impose greater obligations to Internet Intermediaries when dealing with allegedly hateful content. The ECtHR opened the door for this stricter approach when it ruled that Internet Intermediaries can be held liable if they fail to expeditiously delete content that amounts to hate speech and is hosted on their platform.[10] In addition, Germany adopted a new law which imposed an extremely strict obligation to Internet Intermediaries to take down harmful online content within 24 hours; failure to do so would result in paying huge fines. By imposing strict liability on Internet intermediaries, making them obliged to monitor and remove hate speech content, Germany became

[9] EU Commission, 'Code of Conduct on countering illegal hate speech online: First results on implementation'(2016) <http://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupDetailDoc&id=29738&no=1> accessed 12 November 2018

[10] *Delfi AS v Estonia* App no 64569/09 (ECtHR, 16 June 2015), para 159

one of the leading countries in addressing hate speech.[11] On the other hand, there has been some opposite understandings of Internet Intermediaries role and liability model. While in Europe countries are starting to shift towards strict liability model, in India, the Supreme Court ruled that Internet Intermediaries must be provided with broad immunity in order for freedom of expression to be protected.[12] Thus, they will be only held liable in the situation where they do not abide by court or governmental order.[13]

Although the aim of introducing Internet Intermediaries liability is countering online hate speech more properly, this solution can cause certain problems. Prescribing Internet Intermediaries' liability as a way to counter online hate speech requires taking into account three main players and their rights, which have to be properly balanced. Those rights are i) Intermediaries' rights (such as right to conduct business and freedom of expression), ii) rights of alleged victim (such as right not to be discriminated and subjected to violence and hatred), and iii) rights of Internet users who are allegedly inciting violence and hatred (such as right to freedom of expression and right to privacy).[14]

Many concerns arise from prescribing strict liability for Internet Intermediaries. First, determining whether certain content indeed represents hate speech can be extremely hard even for courts. Internet Intermediaries as private companies cannot be expected to play the role of judges, since they do not have proper legal knowledge which would enable them to conclude which content would indeed result in harm. For establishing the unlawfulness of posts legal expertise is

[11] Melissa Eddy and Mark Scott, 'Delete Hate Speech or Pay Up, Germany Tells Social Media Companies' *New York Times* (New York, 30 June 2017) <https://www.nytimes.com/2017/06/30/business/germany-facebook-google-twitter.html> accessed 14 October 2017

[12] *Shreya Singhal vs Union of India* AIR 2015 SC 1523

[13] Suvidutt M. Sundaram and Aditya Tomer, 'Cyberhate in India – Regulation and Intermediary Liability' (2017) 4 (3) International Journal of Law and Legal Jurisprudence Studies 143, 148-149

[14] Christina Angelopoulos and Stijn Smet, 'Notice-and-fair-balance: how to reach a compromise between fundamental rights in European intermediary liability' (2016) 8(2) Journal of Media Law 266, 267

required.[15] Secondly, imposing strict liability to Internet Intermediaries can lead to censorship.[16] Afraid that they could face punishment, Internet Intermediaries could, out of caution, start deleting legal content and suppress Internet users' freedom of expression. As a result, that could have a chilling effect on free speech, so people could start avoiding commenting since they would fear that their comments could be deleted by Internet Intermediaries.[17] Furthermore, strict liability regime may also trigger additional consequences, such as lack of protection for third-party rights, undue burden for Intermediaries, etc.

When discussing all the consequences that online hate speech can have and while trying to find a legal solution for countering it, the role and the nature of the Internet and Internet Intermediaries cannot be disregarded. It is true that Internet Intermediaries are seen as well situated to act against hate speech because they have direct access to harmful material which they host. But, is that enough to justify their increased liability? Thus, the core question of this research will be: Whether imposing strict liability to internet intermediaries is the proper way of combating online hate speech?

In order to answer the question, this research will offer comparative analysis of domestic and regional international documents in force, which prescribe different mechanisms in online hate speech combat and victims' rights protection. It will concentrate primarily on solutions given by the ECtHR, Germany and India. The ECtHR and Germany impose strict liability to Internet Intermediaries requiring from them to act expeditiously and delete hateful content they host on

[15] Office of the Special Rapporteur for Freedom of Expression of the Inter-American Commission on Human Rights, 'Freedom of Expression and the Internet' (2013), para 99

[16] UNHRC, 'Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression Frank La Rue' (2011) UN Doc A/HRC/17/27

[17] Cynthia Wong and James X. Dempsey, *Mapping Digital Media: The Media and Liability for Content on the Internet* (Open Society Foundations, 2011), 10

their platform, which would be possible only by monitoring the content. In contrast, India offers a broad immunity to Intermediaries, obliging them to act only upon court or government order. In addition, the thesis will include relevant international reports and recommendations.

The thesis will proceed in three chapters. First chapter is dedicated to defining hate speech in general terms and pointing out the specific characteristic of online hate speech, as well as all the possible challenges it entails. Therefore, this chapter will first look into different hate speech definitions set down in relevant international documents such as ICCPR, ICERD and CoE Committee of Ministers Recommendation No. R (97) 20 in order to set down a common ground and starting point definition for further thesis analysis. Once it is made clear what hate speech actually entails, the thesis will proceed with singling out all the relevant characteristics of online hate speech, given by different authors and critically assessing each one of them.

From that point, the thesis will proceed with its second Chapter, where the main topic will be establishment of Internet Intermediaries' liability, as one of the solutions for countering online hate speech. The second Chapter will start with defining and distinguishing different types of Internet Intermediaries, putting emphasis on the ones who are hosting third-party content, in other words, ones who are not modifying in any way the alleged hateful third-party content they host, such as internet news portals, social media networks, etc. Then, it will proceed with explaining different models of internet intermediaries' liability adopted worldwide, in order to provide the reader with clear picture of all the possible solutions with regards to this topic. It will elaborate on i) *broad immunity model*, which imposes obligation to Internet Intermediaries to take down illegal content only after obtaining a court order; ii) *safe harbor regime* which imposes obligation to Internet Intermediaries to take down content when they get informed about it by private individuals

(potential victims) or in other way and iii) *strict liability regime* which imposes obligation to Internet Intermediaries to monitor content they host and delete it expeditiously.

From that point, the thesis will turn to explaining the trend that has been evolving among certain jurisdiction with regards to imposing strict obligation to Internet Intermediaries who host third-party content to expeditiously delete allegedly hateful comment. First, it will focus on recent case-law of the ECtHR and its *Delfi AS* judgement where the Court ruled that Internet Intermediaries can be held liable if they fail to expeditiously delete hateful content hosted on their platform.[18] By delivering this judgement, the Court set standards for countering online hate speech through help of Internet Intermediaries, by imposing them stricter obligations, in contrast to other types of illegal content, where such a strict obligation is not requested. The thesis will then move to Germany, where a new law – Network Enforcement Act – regarding Internet Intermediaries liability was adopted. A deep analysis of this law's provisions will be provided. Network Enforcement Act imposes an extremely strict obligation to Internet Intermediaries to take down allegedly hateful online content they host on their platforms within 24 hours. It is "revolutionary act" since it includes not only Intermediaries such as news portals, but also social media platforms, which goes far beyond ECtHR stance. Under the new German law, failure of the social media to take down allegedly hateful online comment hosted on its platform will result into punishment, which consists of huge fines reaching even 50 million euros. By imposing strict liability on Internet intermediaries, making them obliged to monitor and remove hate speech content, Germany became one of the countries with the strictest policy when it comes to countering online hate speech. The thesis will then move on explaining the situation in India after Supreme Court's delivering of a

[18] *Delfi AS* (n 10)

landmark decision[19], which declared unconstitutional Section 66A of the Information Technology Act 2000 prescribing punishments for offences constituted of sending "offensive messages" through communication service[20] which can be regarded as a form of hate speech and by interpreting Section 79 regarding the liability of Internet Intermediaries, stating that they could only be held liable if they fail to obey court or government order.

Thesis will then move on into addressing the negative impact of imposing strict liability model to Internet Intermediaries when dealing with online hate speech they host. This will be done in the final Chapter as an overall assessment of all previously stated. It will single down all the critics, starting from increased possibility of censorship which can be seen as an unavoidable result of imposing an obligation to private entities, such as Internet Intermediaries, to monitor and take down the content that they think it amounts to hate speech. In addition, it will tackle the issue of Internet Intermediaries' lack of legal expertise, which is required in order to properly address whether an allegedly unlawful content indeed represents hate speech and should be strike down or not. Apart from these implications, it will list additional ones, and in the end, it will finish with naming possible reasons that authorities may have to introduce such a regime which is detrimental to core human rights.

The thesis will conclude with emphasizing the non-applicability and detrimental nature of strict liability regime for Internet Intermediaries who host content which may be deemed as hate speech. Thus, it will endorse broad immunity model, which will not hinder neither the importance of the protection of online hate speech victims, nor the other rights at stake.

[19] *Shreya Singhal* (n 12)
[20] Information Technology Act 2000, Gazette of India, Sec 66A

CHAPTER I HATE SPEECH ONLINE

A. Introduction

This chapter will start with listing down different definitions on hate speech adopted in international and regional documents and by national legislature, including the solutions given in ECtHR, Germany and India, as jurisdictions in this thesis's focus. By doing so, it will raise main questions about problems which are encountered when defining hate speech, connected mainly with interpreting terms such as "incitement", "advocacy" etc., and it will tackle the issue of difficulty in establishing whether certain expressions constitute "hate speech" and, thus, should be prohibited. In addition, this part of the chapter will focus as well on the definitions and recommendations given by different authors and it will list important elements common to hate speech definitions. From that point, it will focus on online hate speech by listing down definitions given by regional bodies and different Internet Intermediaries, as they are seen as key players in online hate speech suppression, since they can host content which may be deemed as hate speech. Further on, it will list down main characteristics of online hate speech that make it different from offline one and it will list all the challenges that law makers can encounter. Bearing in mind all the positive sides of Internet for freedom of expression, this Chapter will, as well, briefly address this part, in order to provide the whole picture of the problem, and by doing so, it will lead into the next chapter's discussion about Internet Intermediaries and their presumed role in combating online hate speech.

1. Defining hate speech

Freedom of expression represents a core stone of a democratic society[21], a right explicitly guaranteed under many international documents.[22] Suppressing and silencing speech without giving people the freedom to speak their mind can endanger stability and democracy of a society.[23] Broad limitations on free speech can be used not only as a tool for punishing political opponents, but they can also result in suppressing the truth.[24] Thus, all the limitations on free speech must be carefully construed, in order to avoid misuse.

One of the reasons for limiting freedom of expression is protection of the rights of others, but as well respect for human dignity and equality, which, in the same manner as free speech, represents one of the basic principles on which our society is built. For this reason, international and national communities introduced the notion of "hate speech"[25], a type of expression which by inciting hatred and violence causes harm to members of certain groups towards whom it is directed. This means that not all people would be protected from hate speech, but only the ones who are members of certain groups. The idea is based on the stance that some people are in more "vulnerable" position due to the higher possibility to become a victim of different offences because of personal characteristics they possess. Thus, addressing those groups and its members as "vulnerable", different jurisdictions prescribe different scope of groups which may enjoy protection from hate

[21] *Perinçek v Switzerland* App no 27510/08 (ECtHR, 15 October 2015), para 196 (i)

[22] Universal Declaration of Human Rights (adopted 10 December 1948) UNGA Res 217 A(III) (UDHR) art 19; International Covenant on Civil and Political Rights (adopted 16 December 1966, entered into force 23 March 1976) 999 UNTS 171 (ICCPR) art 19 (2); African Charter on Human and Peoples' Rights (adopted 27 June 1981, entered into force 21 October 1986) (1982) 21 ILM 58 (AfCHR) art 9 (2); American Convention on Human Rights (adopted 22 November 1969, entered into force 18 July 1978) 1144 UNTS 123 (ACHR) art 13 (1); European Convention on Human Rights (adopted 4 November 1950, entered into force 3 September 1953) (ECHR) art 10 (1)

[23] *Whitney v California* 274 US 357 (1927) concurring opinion of judge Brandeis

[24] John Stuart Mill, *On Liberty* (first published 1859, Batoche Books 2001); *Abrams v US* 250 US 616 (1919) dissenting opinion of judge Holmes

[25] The term "hate speech" in Black's Law Dictionary is defined as "a speech that carries no meaning other than the expression of hatred for some group, such as a particular race, especially in circumstances which the communication is likely to provoke violence". See Bryan A Garner (ed.), *Black's Law Dictionary* (8rd edn, The West Group 2014)

speech depending on the historical circumstances and social context in the given state. Nowadays, protected groups are usually LGBT people, women, people with disabilities, elderly, children, people of different ethnic or national origin, people of different race, etc. This may provoke different legitimate questions, such as how far the legislative solutions can go in giving special protection to different groups based on their certain personal characteristics. Nevertheless, one thing is certain. Hate speech can be especially harmful when it is addressed towards the groups that were historically oppressed, as it is the case with racism.

This at the same time raises the question of potential victims. By how it is usually defined, hate speech could be directed towards a group of people without the need to be directed towards a specific individual victim. In other words, a person, who is a member of a certain group protected under national hate speech provisions e.g. a person of black race, may claim to be a victim of a speech which incited to hatred towards all the members of black race in general, and not him directly.

Nowadays, there is no universally accepted definition of hate speech and the ones that do exist, are often being criticized for not being precise enough. One of the most important international human rights documents which is often being referred to when addressing the notion of "hate speech" is ICCPR and its Articles 19 and 20.[26] It must be noted that ICCPR, does not explicitly mention the term hate speech, but it offers grounds – such as e.g. "the rights of others", "public order", etc. – for prohibiting speech that could be regarded as "hate speech". To be more precise, in its Article 19(3), ICCPR sets down the three-tier test for limiting the right to freedom of expression, which could be called upon when an expression is deemed to be limited since it

[26] E.g. see *R v Keegstra* 3 S C R 697 (1990)

amounts to hate speech. In addition, in its Article 20(2) – which is regarded as fully compatible with Article 19[27] – ICCPR prescribes prohibition of "*any advocacy* of national, racial or religious hatred that *constitutes incitement to discrimination, hostility or violence*". This prohibition is seen as one of the forms of hate speech, and usually, hate speech is therefore, primarily, defined as incitement to hatred and violence. In that manner, UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, as well, uses term "incitement to hate" when referring to hate speech.[28]

Another important international document for understanding the notion of the term hate speech and what it actually entails is ICERD. ICERD as well, although not explicitly, forbids certain forms of hate speech, namely the one addressed to race and ethnicity. In its Article 4 (a) it provides prohibition of "*dissemination* of ideas based on racial superiority or hatred, *incitement* to racial discrimination, as well as all *acts of violence* or *incitement to such acts* against any race or group of persons of another color or ethnic origin".[29] Thus, ICERD uses terms such as "dissemination" and "incitement to discrimination and violence". In addition, it imposes on States obligation to criminalize such behaviors, which represents more restrictive obligation than the one set in the ICCPR. This view is justified due to the fact that racist speech represents one of the most severe forms of the harmful expressions addressed towards the group which is historically being oppressed.[30]

[27] UNHRC, 'General Comment No. 11: Prohibition of propaganda for war and inciting national, racial or religious hatred (Art. 20)' (1983)
[28] UNGA, 'Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression' (2012) UN Doc A/67/357, chap III
[29] International Convention on the Elimination of All Forms of Racial Discrimination (adopted 21 December 1965, entered into force 4 January 1969) (ICERD) art 4 (a)
[30] Such view originates from the Holocaust period and represents the solution aimed at stopping the crimes, such as the one committed during the Second World War, to re-occur.

Another important international document related to hate speech is Convention on the Prevention and Punishment of the Crime of Genocide which prohibits "direct and public incitement to commit genocide"[31]. This Convention obliges countries to criminalize such behavior, since, this form of expression is, like the previous mentioned, one of the most severe forms of the harmful expressions. Genocide Convention uses only term "incitement" for prescribing the punishable act, adding that such incitement must be direct and public, in that manner excluding the prohibition of indirect incitement. This shows the variety of the approaches adopted by different international documents and different forms of what can be possibly referred to as hate speech.

Bearing in mind these international documents, NGO Article 19 in interpreting the provisions of international documents makes difference among certain types of hate speech depending on the severity of the speech.[32] First, "incitement to genocide" set forth in Convention on the Prevention and Punishment of the Crime of Genocide and "advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence" set forth in Art 20(2) of the ICCPR must be prohibited as the most severe forms of "hate speech". When it comes to expression which endangers reputation and the rights of others, national security, public order, moral or health, this type of expression may be restricted if it meets the criteria from Art. 19(3) of the ICCPR, where a Court shall provide a proper balancing through apply proportionality test and take into account all relevant circumstances of the case. At the end, intolerant speech must be protected, since the speech that can offend should enjoy protection to certain limit. Thus, it must be stressed that freedom of expression does not apply only to information or ideas that are favorably received

[31] Convention on the Prevention and Punishment of the Crime of Genocide (adopted 9 December 1948, entered into force 12 January 1951) art 3 (c)

[32] Article 19, 'Hate speech' Explained – A toolkit' (2015) p 19-23 <https://www.article19.org/data/files/medialibrary/38231/'Hate-Speech'-Explained---A-Toolkit-%282015-Edition%29.pdf> accessed 4 February 2018

but also to those that shock, offend and disturb the State or any sector of the population.[33] All of this makes it hard to draw a line between lawful and unlawful speech even for Courts.

When it comes to addressing the hate speech issue on regional level, certain differences can be encountered here, as well. Council of Europe and Committee of Ministers were one of the first bodies to give a definition of hate speech as such, stating that it covers "all forms of expression which spread, incite, promote or justify racial hatred, xenophobia, anti-Semitism or other forms of hatred based on intolerance, including: intolerance expressed by aggressive nationalism and ethnocentrism, discrimination and hostility against minorities, migrants and people of immigrant origin".[34] As it can be seen, Committee of Minister's definition includes terms as "spreading", "inciting", "promoting", "justifying" which all represent different level of seriousness and some of them are more restrictive to freedom of expression. At the same time, within the Council of Europe, the ECtHR decides hate speech cases on the basis of Article 10 of the European Convention on Human Rights, which sets general limitations to freedom of expression in the form of three-tier test.[35] The ECtHR tends to take different approaches when deciding these cases. In cases of holocaust denial[36] and "serious" racist hate speech[37], the Court usually declares applications inadmissible relying on Article 17, finding them not appropriate to enjoy the protection under the Convention since they represent abuse of rights, while in other hate speech

[33] *Handyside v the United Kingdom* App no 5493/72 (ECtHR, 7 December 1976), para 49; *Mouvement raëlien suisse v Switzerland* App no 16354/06 (ECtHR, 13 July 2012), para 48; *Steel and Morris v the United Kingdom* App no 68416/01 (ECtHR, 15 February 2005), para 87; *Good v Botswana* (2010) AHRLR 43 (ACmHPR 2010), para 198; *"The Last Temptation of Christ" (Olmedo-Bustos et al) v Chile* Inter-American Court of Human Rights Series C No 73 (5 February 2001), para 69

[34] CoE, 'Recommendation No. R (97) 20 of the Committee of Ministers to Member States on Hate Speech' (1997) R (97) 20

[35] European Convention on Human Rights (n 22), art 10 (2)

[36] *Garaudy v France* App no 65831/01 (ECtHR, 24 June 2003)

[37] *Norwood v UK* App no 23131/03 (ECtHR, 16 November 2004)

cases it applies Article 10 and decides the cases on merits applying three-tier test.[38] In assessing the limit between freedom of expression and hate speech, the Court tries to provide proper balance between competing rights.[39] In the ECtHR view, hate speech, seen as expression that "promotes or justifies hatred, violence, xenophobia or another form of intolerance"[40] cannot be protected under the European Convention on Human Rights. However, sometimes, even the ECtHR has problems in deciding whether certain speech should be regarded as hate speech or not, due to the thin line between hate speech and freedom of expression. As it will be shown in the next Chapter of this thesis, some rulings on hate speech provoke many concerns, as the ruling in the *Delfi AS* case did, where the Court characterized online comments on an online news article as hate speech, although they did not incite any violence nor discrimination.

When it comes to American region, apart from general limitation imposed on free speech, American Convention on Human Rights in its Art. 13 (5) similarly as Art. 20 (2) of the ICCPR uses terms such as "advocacy" and "incitement" and requires "any advocacy of national, racial, or religious hatred that constitute incitements to lawless violence or to any other similar illegal action" to be punishable, adding that such an action is prohibited if it is directed towards "any person or group of persons on any grounds including those of race, color, religion, language, or national origin".[41] As noticeable, the scope of protected groups here is very broad.

To make things even more complicated, national legislations as well differ when prescribing hate speech. Since this thesis will, apart from already mentioned ECtHR's jurisdiction, compare as well

[38] See e.g. *Balsytė-Lideikienė v Lithuania* App no 72596/01 (ECtHR, 4 November 2008); *Féret v Belgium* App no 15615/07 (ECtHR, 16 July 2009); *Leroy v France* App no 36109/03 (ECtHR, 2 October 2008); *Jersild v Denmark* App no 15890/89 (ECtHR, 23 September 1994)
[39] *Perinçek* (n 21)
[40] ibid, para 230
[41] American Convention on Human Rights (n 22), art 13(5)

legislative solutions in Germany and India, a brief overview of their solution in prescribing hate speech is put forward. In that manner, it is worth mentioning that Germany in its Criminal Code criminalizes hate speech through the offence "incitement to hatred". Thus, in Germany "incitement to hatred" and "call for violent or arbitrary measures" which are aimed "against a national, racial, religious or ethnic group or an individual belonging to one of the mentioned groups" are punishable under criminal law.[42] This is a logical result, since in Germany respect for human dignity represents a principle usually overriding other basic rights. Germany, as a country that went through holocaust, seems as particularly sensitive to the issues of hate speech, as it will be shown through the analysis of their newly enacted law dealing with online hate speech.

On the other hand, in India, the regulation of hate speech has its basis in Article 19(2) of the Indian Constitution which sets restrictions on freedom of expression, especially in the interests of the sovereignty and integrity of India, public order and an incitement to an offence.[43] The Constitution does not explicitly mention hate speech; however, its judiciary has recognized it and defined it. In the case *Pravasi Bhalai Sangathan v. Union of India and Ors.*, the Supreme Court of India defines hate speech as "*an effort to marginalize individuals based on their membership in a group*",[44] which has an impact both on individuals and on society and may lead to discrimination, segregation, deportation, violence or even genocide.[45] However, "the effect of the words must be judged from the standards of reasonable, strong-minded, firm and courageous men, and not those

[42] Criminal Code in the version promulgated on 13 November 1998, Federal Law Gazette [Bundesgesetzblatt] I p 3322, last amended by Article 1 of the Law of 24 September 2013, Federal Law Gazette I p 3671 and with the text of Article 6(18) of the Law of 10 October 2013, Federal Law Gazette I p 3799, art 130(1)

[43] Gargi Chakrabarti and Saahil Dama, 'Intermediary Liability and Hate Speech' <https://www.law.uw.edu/media/1395/india-intermediary-liability-of-isps-hate-speech.pdf> accessed 15 October 2018

[44] *Pravasi Bhalai Sangathan v Union of India and Ors* (2014) SC 1591, para 7

[45] ibid, para 7

of weak and vacillating minds, nor of those who scent danger in every hostile point of view."[46] In that manner, speech that merely creates feelings of disaffection or enmity in people is protected under the Constitution.[47]

Additionally, hate speech is punishable under different provisions of Indian Penal Code, such as Section 153A which criminalizes "*promotion of enmity* between different groups on grounds of religion, race, place of birth, residence, language, etc., and doing acts prejudicial to maintenance of harmony"[48] and Section 295A which criminalizes "deliberate and malicious acts, intended to outrage religious feelings of any class by insulting its religion or religious beliefs"[49]. While deciding upon the constitutionality of the Section 295A, the Supreme Court emphasized that only an aggravated and deliberate form of insult to religion is to be considered criminalized under this Section and not in breach of freedom of expression.[50]

As seen, most of the legislative definitions include terms such as "incitement", "advocacy" "promotion", "propaganda", etc. and all of them provide different level of protection, often unclear. Therefore, one of the problems with defining hate speech and the way it is prescribed in law comes with the lack of setting clear standards. Even same terms may be interpreted differently in different jurisdictions. In that manner, in some jurisdictions, such as US, in order to be prohibited, speech must be likely to lead to imminent violence, requiring "incitement" to be interpreted in that manner. On the other hand, in most European jurisdictions, the link between the violence and the speech does not need to be direct.

[46] *Ramesh v Union of India* AIR 1988 SC 775
[47] *Kedar Nath Singh v State of Bihar* 1962 Supp (2) SCR 769
[48] The Indian Penal Code, Gazette of India Sec 153A
[49] ibid, Sec 295A
[50] *Ramji Lal v State of Uttar Pradesh* 1957 AIR 620

It is hard to establish what type of speech is actually "advocacy… of hatred that constitute incitement"[51]. Different countries and different bodies have their own way of interpretation, which further leads to inconsistency and unpredictability with defining "hate speech". Too broadly set, terms like "advocacy", "promotion", even "incitement", etc. can lead to misuse. These terms can be used especially as a suitable tool for silencing the speech of political opponents,[52] which is, for example, especially present in India.[53] Thus, there have been initiatives aimed at defining all of the problematic terms.[54] With the aim to provide universal approach to it, organization Article 19 defined "incitement" as "statements made about national, racial or religious groups that create an *imminent risk* of discrimination, hostility or violence against persons belonging to those groups"[55], while "advocacy" was defined as "intentionally promoting hatred publicly towards the target group"[56]. However, this does not resolve many problems that are encountered when defining hate speech and when provisions are actually applied in practice.

All of the so far mentioned opens some of the crucial questions. What are the groups of people that should be protected from hate speech; what is the required level and type of harm that should be caused in order to claim violation; whether the harm should be imminent or likely to happen; whether is it enough that the speech is directed towards the group in the abstract sense or there should be a directly affected individual victim, should intent be present, and similar.[57] These issues make it hard to establish the existence of hate speech and set the limit between freedom of

[51] International Covenant on Civil and Political Rights (n 22), art 20 (2)
[52] UNESCO, 'World Trends in Freedom of Expression and Media Development' (2015) p 27 <http://unesdoc.unesco.org/images/0023/002349/234933e.pdf> accessed 9 November 2018
[53] Gargi Chakrabarti and Saahil Dama (n 43)
[54] Article 19, 'The Camden Principles on Freedom of Expression and Equality' (2009) prin 12 <https://www.article19.org/data/files/pdfs/standards/the-camden-principles-on-freedom-of-expression-and-equality.pdf> accessed 31 January 2018
[55] ibid, prin 12.1. (iii)
[56] ibid, prin 12.1. (ii)
[57] Article 19 (n 32), p 10-11

expression and hate speech. Even the courts have trouble in assessing those questions and drawing a clear line between freedom of expression and hate speech, as already mentioned. Thus, when assessing potential hate speech cases, the active role of the courts is required, which presupposes taking into account all the necessary circumstances of each case in order to determine whether certain expression is severe enough to justify limiting freedom of expression. The factors usually taken into account by courts are content and form, context (including audience, the place and time, etc.), the speaker, severity of the expression, extent, including the media used for expressing and size of the audience, imminence or likelihood of violence to happen.[58] Only thorough and proper examination of all the factors can lead to the delivery of a good decision. This is very important in the light of the present thesis topic regarding the direction in which cybercrime regulations are moving by imposing obligations to Internet Intermediaries to monitor content and self-initiatively decide which content could be regarded as hate speech. It poses the question whether private entities such as those can recognize the thin line between hate speech and freedom of expression.

When it comes to scholars and their view on hate speech, some authors especially emphasize that both the speeches which can have direct or indirect harmful effect on the victim – meaning that they can be directly aimed at victim or can be a call for others to act against certain people – should be prohibited.[59] Some authors state that when defining hate speech, law should take into account only incitement to xenophobic hatred, without naming different protected groups, since xenophobic hatred would include hate against all the potential groups, and specific nature of victim

[58] Article 19, 'Towards an interpretation of article 20 of the ICCPR: Thresholds for the prohibition of incitement to hatred' (2010), p 10-17 <http://www.ohchr.org/Documents/Issues/Expression/ICCPR/Vienna/CRP7Callamard.pdf> accessed 1 February 2018; UNGA 'Rabat Plan of Action on the prohibition of advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence' (2012) A/HRC/22/17/Add.4, para 29

[59] Susan Benesch, 'Dangerous Speech: A Proposal to Prevent Group Violence' (2012) World Policy Institute <http://www.worldpolicy.org/sites/default/files/Dangerous%20Speech%20Guidelines%20Benesch%20January%202012.pdf> accessed 2 February 2018

group, such as race, color, etc., should be assessed as aggravating circumstances.[60] Others think that protected grounds from hate speech should be the same as the ones listed in non-discrimination provisions.[61] There are also authors, who tend to draw difference between different level of harmful speech, stating that a speech that is more likely to lead to violence towards certain group, should be regarded as "dangerous speech".[62] Some emphasize the fearful effect of harmful speech, meaning that the aim of the what is called hate speech is to cause fear among protected groups of people, and therefore instead of using "hate speech" for a term, term "fear speech" would suit better.[63]

At the end, the only conclusion that can be drawn is that there is no single universally accepted definition, and even, when certain bits are provided, the interpretation differs a lot, which makes a lot of confusion for all important legal actors. However, taking into account all definitions given on international, regional and national level and by different scholars, a tentative conclusion may be drawn that a certain expression may be regarded as hate speech if it consists of two main elements – first, that expression constitutes incitement of hatred and violence and second, that such incitement is directed towards identifiable vulnerable group, in other words towards the collective as such.[64]

However, it remains an open question whether even this is a good way to define hate speech. In that manner, some authors have pointed out that hate speech regulations put feelings in front of rights, due to the people's nature to try to fight and silence those who spread hate and insults

[60] Chris Reed, 'The Challenge of Hate Speech Online' (2009) 18 (2) Info & Comm Tech L 79, 82
[61] Article 19 (n 32), p 15
[62] Susan Benesch (n 59)
[63] Antoine Buyse, 'Words of Violence: "Fear Speech," or How Violent Conflict Escalation Relates to the Freedom of Expression' (2014) 36 (4) Human Rights Quarterly 779
[64] James Hawdon, Atte Oksanen and Pekka Räsänen, 'Exposure to Online Hate in Four Nations: A Cross-National Consideration' (2017) 38 (3) Deviant Behavior 254, 254-255

towards them, thus making those regulations very dangerous and subject to abuse.[65] However, these regulations do exist and hate speech as such is criminalized in many jurisdictions.

2. Defining hate speech online: characteristics and challenges

Raising of Internet technology has set many challenges for a mankind. For many, Internet has been seen as a tool for performing different types of cybercrimes, including hate speech.[66] Many of perpetrators of hate speech use internet with aim to encourage others to adopt their way of thinking and join their ideology of spreading hatred towards specific "vulnerable" groups.[67]

All people must enjoy protection of their rights, not only in offline world, but as well on Internet.[68] The same applies to protection of rights and dignity of people which can be endangered by online hate speech. Although it resembles the offline one, many authors emphasize that online hate speech entails certain characteristics and its suppression often encounters many challenges which are related to the nature of Internet. However, the question is to what extent does online hate speech differ from offline one.

Emphasizing the specifies and dangers which online hate speech can have, in order to address separately this issue, Council of Europe adopted an important regional document dealing with it. Additional Protocol to the Convention on Cybercrime, concerning the criminalization of acts of a racist and xenophobic nature committed through computer systems deals especially with racist and xenophobic type of hate speech. It prohibits dissemination on racist and xenophobic material which "advocates, promotes or incites hatred, discrimination or violence against any individual or

[65] George Packer, 'Mute Button' *The New Yorker* (New York, 13 April 2015) <https://www.newyorker.com/magazine/2015/04/13/mute-button> accessed 7 November 2018
[66] James Banks, 'European Regulation of Cross-Border Hate Speech in Cyberspace: The Limits of Legislation' (2011) 19 European Journal of Crime, Criminal Law and Criminal Justice 1, 3-4
[67] James Hawdon, Atte Oksanen and Pekka Räsänen (n 64), 255
[68] UNGA, 'The promotion, protection and enjoyment of human rights on the Internet' (2012) A/HRC/RES/20/8, para 1

group…based on race, religion, color, descent or national or ethnic origin".[69] The offence is committed through computer system, which represents an important characteristic of the offence. In addition, the Additional Protocol also prohibits "racist and xenophobic motivated threat and insult"[70] and "denial, gross minimization, approval or justification of genocide or crimes against humanity"[71]. As an obligation taken by signing it, Member States should criminalize all of these offences in their national laws.

The general problem when it comes to online hate speech is the fact that the states are not the only ones who define hate speech in their law, but as well, Internet Intermediaries, as the ones who can host different type of content, play a crucial role.[72] In their Terms of Use and Service they define the type of content they will not tolerate on their platforms. Thus, for example, Facebook in its Terms of Use explicitly says that users should not post content which represents hate speech or incites violence[73] and violates others' people rights or law[74]; otherwise, their content may be removed.[75] As well, it states that organized hate groups will not be allowed, as well as content which represents support for such group and their violent acts.[76] On the other hand, Twitter, does not allow hateful content directed towards certain group explicitly stated in their Hateful Conduct Policy.[77] Twitter, opposite of Facebook, provides closed list of grounds when it comes to hateful

[69] Additional Protocol to the Convention on Cybercrime, concerning the criminalization of acts of a racist and xenophobic nature committed through computer systems (adopted 28 January 2003, entered into force 01 March 2006) ETS No 189, art 3
[70] ibid, art 4-5
[71] ibid, art 6
[72] More about Internet Intermediaries and their role will be elaborated in the next Chapter.
[73] Facebook, 'Statement of Rights and Responsibilities' para. 3 (7) <https://www.facebook.com/legal/terms> accessed 1 February 2018
[74] ibid, para 5 (1)
[75] ibid, para 5 (2)
[76] Facebook, 'Community standards: Helping to keep you safe' <https://www.facebook.com/communitystandards#dangerous-organizations> accessed 1 February 2018
[77] Twitter, 'Hateful conduct policy' <https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy> accessed 1 February 2018

conduct and hate speech, such as "race, ethnicity, national origin, sexual orientation, gender, gender identity, religious affiliation, age, disability, or disease", and it forbids "promotion of violence against or directly attacking or threatening" or "inciting harm" towards people that fall within one of the mentioned category.[78] It lists certain behaviors which could be regarded as harmful and, thus, removed, such as "incitement to fear among protected group", "threat of violence", "wishes for physical harm, death or decease", "references to mass murder, violent events, or specific means of violence in which such groups have been the primary victims", etc.[79] As well, it emphasizes that context of the content must be taken into account when assessing its unlawfulness by the Twitter team.[80] When it comes to YouTube, its Community Guidelines sets hate speech definition which in their views represent any content that "promotes violence against or has the primary purpose of inciting hatred against individuals or groups based on certain attributes" with naming a number of grounds.[81] Interestingly, YouTube, giving an example, states that a line can be drawn between hate speech and freedom of expression, and that a content will not represent hate speech unless its only purpose is to incite hatred toward a person or a group based on its characteristic or to promote violence against them.[82] But, how can any of these platforms decide what is inciting hatred and what is not? By giving their own view of what hate speech is, they can arbitrarily remove content which they see as harmful, which can endanger free speech, as it will be further elaborated in the upcoming chapters. Although Internet Intermediaries should be subjected to the national laws and their defining of hate speech, this can be additionally complicated with the fact that most of them operate in many jurisdictions, and different

[78] ibid
[79] ibid
[80] ibid
[81] YouTube, 'Hate speech' <https://support.google.com/youtube/answer/2801939?hl=en> accessed 1 February 2018
[82] ibid

jurisdictions can differently approach the issue of hate speech. This leads to lack of possibility to effectively apply domestic legal norms,[83] since there is not a clear line which national law would apply; thus, Internet Intermediaries provide their definitions of hate speech and implement their own interpretation of it.

The problem with characterizing certain online content as hate speech comes from blurring relation between offline violence and online hate speech,[84] since sometimes the lack of the increased likelihood of violence linked to harmful content can lead to not defining certain content as hate speech. This is usually related with the widespread stance that people acting online in most cases would not make any step into the offline world against a certain person. Whoever assesses whether the online content constitutes hate speech, must consider all the relevant factors already mentioned, such as context, content, etc. Thus, only a proper judicial examination of all the facts of the case can result in establishing whether an online content calling for violence against a certain person would have the same effect as the face-to-face.[85]

2.1. Anonymity of perpetrators

There are many characteristics related to online hate speech, which are raised by many authors as important for distinctions between online hate speech and offline one. First important characteristic find in the literature is the possibility of perpetrators to act and remain anonymous online. Many authors claim that online hate speech is more present than offline one, since perpetrators feel more comfortable in expressing their calls for violence and hate towards people, because there are aware of the possibility to hide their identities and not face punishment.[86] That being the case, victims

[83] UNGA (n 28), para 32
[84] UNESCO (n 52), p 26
[85] *Delfi AS* (n 10) dissenting opinion of judges Sajo and Tsotsoria, para 14
[86] Interview with Drew Boyd, Director of Operations, The Sentinel Project for Genocide Prevention (24 October 2014)

may never find out who is the person inciting hatred towards them[87] and may never be remedied, which can inflict even more harm.[88] Although this resembles the reality to some extent, however, it should be noted that, at the same time, anonymity is an important factor which can influence people to feel free to criticize and express their opinion online, without the fear of being prosecuted for their views, which is as an important element of democratic society. As such, anonymity can be seen as an important tool for people to contribute to public debate though freely expressing their ideas and should be regarded as a value.[89] Thus, this shall be borne in mind when adopting laws regulating online hate speech. Additionally, it is important to note that anonymity can be revealed as well. The technology has developed to such extent that revealing of identity has become much easier. In that manner, for example, using different tools for revealing IP addresses can lead to disclosure of anonymous user's identity which can then lead to prosecution of the anonymous perpetrators. Only a person with high technological skills and expertise may be capable of blurring his or her track. Moreover, even in the real world, perpetrators may incite hatred anonymously. For example, a person may distribute anonymous flyers with messages which call for violence by putting them in people's mails.

Judge Zupančić in his concurring opinion in the ECtHR case *Delfi AS*, which will be the subject of deeper analysis in the third chapter, claimed that online media portals should not be anyhow permitted to allow third-party anonymous comments on their news publication, since such comments tend to be more insulting.[90] However, such view is quite controversial and damaging

[87] Dragos Cucereanu (n 2), 183
[88] Tarlach McGonagle, (n 4) p 29 –30
[89] *Delfi AS* (n 10), para 95
[90] *Delfi AS* (n 10), concurring opinion of judge Zupančić

towards people's right to privacy and freedom of expression. Additionally, as explained, it may lead to decrease of public debate and discourse.

2.2. Easily accessible, visible and itinerant content

Furthermore, it is believed that, due to the Internet characteristics, hate speech has become easily accessible and visible[91]. The main problem with online hate speech – that comes from its boundarylessness and it is tightly related to it – is that the content can be easily reached from any place. For example, a hateful anti-Semitic content written on English news portal can be read by a Jewish person in Hungary or elsewhere. The audience can be vast, especially if content goes viral[92], which is nowadays much probable due to the use of social media. Such content can cause great harm to victims and lead to emphasized sense of embarrassment.[93] However, at the same time, it is important to draw a line between content that goes viral, and the one that does not. The consequences of viral content may be completely different from the one that does not reach any attention. The way Internet is functioning must be borne in mind, and the fact that many of the content may not go viral. Content aimed towards certain group of people, may not be even reaching them, since usually the recipients will be people with similar ideas who search for certain content, since most of the users use Internet for this reason.[94] In that sense, the material is usually not imposed to potential victims, quite opposite, people have to search for it.[95]

[91] UNGA (n 28), para 24, 30
[92] Dragos Cucereanu (n 2), 193
[93] Chara Bakalis, 'Rethinking cyberhate laws' (2018) 27 (1) Information & Communications Technology Law 86, 102
[94] Bibi van Ginkel, 'Incitement to Terrorism: A Matter of Prevention or Repression?' (2011) ICCT Research Paper <http://www.icct.nl/download/file/ICCT-Van-Ginkel-Incitement-To-Terrorism-August-2011.pdf> accessed 2 February 2018
[95] *Féret* (n 38) dissenting opinion of judge Sajó joined by judges Zagrebelsky and Tsotsoria

In addition, online hate speech can be itinerant,[96] meaning that once deleted content can re-appear on the same or different platform. A person writing certain hateful content can write it again, using the same or a different name, with possibility of doing this extremely fast. Even if a person gets blocked for posting hateful content, he or she can open a new account. The same goes with e.g. creating website pages, blogs, or Facebook pages. Once deleted or blocked, they can be again created. Since hateful content can remain online forever, some authors emphasize that it is difficult not to be exposed to the hateful content again on Internet, thus the harm caused by online hate speech is increased.[97] However, although in the offline world, where, upon punishing the perpetrator, the spreading of content is stopped, this may also be temporarily, and hate speech messages may re-appear, since perpetrator can again engage in inciting hatred.

2.3. Court's jurisdiction in cases of online hate speech

Additional characteristic is deeply related to the nature of Internet and its boundarylessness which leads to difficulties in establishing the court's jurisdiction in dealing with cases of online hate speech. Online content which represent hate speech can be published by nationals of one country, directed towards the nationals of second country and hosted on platform which operates in number of different countries. In that situation, it is very hard to establish which Court would have the jurisdiction. Even if a perpetrator is found, due to the limitlessness of the Internet, maybe he or she cannot be brought to justice due to the jurisdiction limitations. But apart from this, this problem is especially important since different jurisdictions differently define hate speech, as already stated. What may be hate speech in one country, may not be in another. For these reasons, even enforcement of judgements in certain countries can be endangered. For example, if a hateful

[96] UNESCO, 'Countering online hate speech' (2015) p 13-14 <http://unesdoc.unesco.org/images/0023/002332/233231e.pdf> accessed 25 November 2018
[97] Chara Bakalis (n 93), 103

content is hosted by an internet platform based in a country where such content could not be regarded as hate speech, and the judgement which obliges a platform to delete such content is rendered by the court in another country where this content represents hate speech, it will be hard to enforce it in country where platform is based.

The UK and French cases may serve as good examples for jurisdiction issue. UK courts in the case of Shepard and Whittle[98] accepted the jurisdiction although the harmful material defendants posted on website based in California, basing such decision on the fact that material was aimed to reach UK public, it was published and edited in the UK and defendants, who were based in UK, could have control it and remove it. However, some authors ask legit question such as what would have happened if the circumstances had been different and included more jurisdictions, for example, that defendants were based in some other country?[99] Another interesting case is French case of *UEJF et LICRA c Yahoo! Inc et Yahoo France*[100], case in which the defendants were allowing their online auction service to be used for selling of Nazi memorabilia which was contrary to the French criminal law. Although defendants claimed that their services were based at US and aimed at US residents and that their right to freedom of expression guaranteed under the First Amendment would unable enforcement of French courts judgement, French courts took the stance that they can have jurisdiction as long as the harmful material can be downloaded in their country, which was possible in the case in question.

2.4. Characteristics related to victim

Additional characteristic related to the online hate speech are concerning the targets of the hate speech. Most of the online hate speech is targeting individuals based on their ethnicity and

[98] Sheppard and Whittle [2010] EWCA Crim 824
[99] Chara Bakalis (n 93), p 96
[100] TGI Paris, 20 novembre 2000, UEJF, LICRA et MRAP (intervenant volontaire) c/ Yahoo! Inc. et Yahoo France

nationality,[101] although other grounds are present as well. Some authors argue that online world can bring a lot of new groups which can be victims of hate speech, so it would be hard to decide whether these new groups qualify as protected and how to broaden the scope.[102] In addition, when it comes to victims of online hate speech, it is believed that the risk of being a victim is increased if a person is more likely to use many different internet services, visits websites where harmful content is more likely to be found, reveals personal information and easily trust people.[103]

2.5. Participation of Internet Intermediaries which host online hate speech

However, maybe one of the most important challenges and characteristics of online hate speech recently developed is the participation of intermediaries in hosting online hate speech and their role in its suppression. Online hate speech brings together three main players, which are the perpetrator, the victim and Internet Intermediary.[104] Although not all the harmful content is hosted by Internet Intermediaries, most of it is. Internet Intermediaries enable exchanges of the hate speech online, by hosting the hateful content. Thus, they are usually invited to cooperate with states in suppressing harmful content. Many questions related to Internet Intermediaries remain to be answered, and one of the most important one is the question of their liability.

3. Freedom of expression and the Internet

At the same time while addressing the question of hate speech online, the role of the Internet in enhancing freedom of expression must be elaborated as an issue non-separable with online hate speech. "The Internet has now become one of the principal means of exercising the right to

[101] Hatebase, 'Most common hate speech' <https://www.hatebase.org/popular> accessed 1 February 2018
[102] Chris Reed (n 60), 79
[103] James Hawdon, Atte Oksanen and Pekka Räsänen (n 64), 256
[104] Christina Angelopoulos and Stijn Smet (n 14), 267

freedom of expression and information, providing as it does essential tools for participation in activities and discussions concerning political issues and issues of general interest."[105]

Rise of the Internet undoubtedly helped fostering freedom of expression. Internet and Internet Intermediaries, as important subjects of this thesis, offered place where people could exchange ideas and engage into fruitful debates, and where public debate and democracy may be obtained.[106] Exchanging ideas and engaging in debates is extremely beneficial for people because this way they form their own opinion. At the same time, Internet and especially Internet Intermediaries offered place for groups regarded as marginalized to share their views and opinions.[107]

Additionally, by virtue of Internet sharing ideas and opinions between people became possible regardless of the distance and time. In that manner, Internet enables people from other continents to discuss different topics, without the necessity to do that in the real time. Namely, Internet Intermediaries have the ability to store messages, which enables users to comment and engage in discussion in any time they want.

Moreover, Internet has ability to bring the information to users.[108] For example, by using search engine, a person can find an unimaginable quantity of the information relevant to his preferences. As well, the hyper-linking tools make easier for a reader to access different information from the content he or she is researching, without the need to look for it in different places – the bare text will lead and enable him or her to extensively explore certain topic.[109] In addition, social media

[105] *Ahmet Yildirim v Turkey* App no 3111/10 (ECtHR, 18 December 2012) para 54
[106] *Delfi AS* (n 10), dissenting opinion of judges Sajo and Tsotsoria, para 6; David Harvey, *Collisions in the Digital Paradigm: Law and Rule-making in the Internet Age* (Hart Publishing 2017), 37-38
[107] Dragos Cucereanu (n 2), 166-168
[108] David Harvey (n 106), 35
[109] ibid, 25-27

platforms such as Facebook and Twitter made it possible for every people using it to send and receive information at the same time.[110]

Freedom of expression is particularly enhanced by the fact that the usage of Internet for sharing ideas is much more cost-effective than TV or other media.[111] Everyone can use Internet for sharing information, without being required to have some special knowledge. In that manner, OSCE has recognized the importance of Internet for freedom of expression, outlining the fact that people are more eager to engage in the exchange of opinions and debates online, since it is easier, more cost-effective and it enables them to remain anonymous.[112]

Additionally, it is worth mentioning that Internet has a different impact than the radio, television or even leaflets, because the opinions are not imposed.[113] The recipients of information are persons with similar ideas[114] who actively search for them.[115] Furthermore, the statements are easily avoidable[116] by not following certain pages, websites, online newspaper, accounts, etc., making the impact of spreading hateful ideas limited. Additionally, the fact that hate speech is spread virtually instead of online, decreases tensions and anger.[117] These are the facts which must be born

[110] Rolf H Weber, 'Challenges for Communications in a Changing Legal Landscape' in Doreen Weisenhaus and Simon NM Young (eds), *Media Law and Policy in the Internet Age* (Hart Publishing, 2017), 167
[111] Dragos Cucereanu (n 2), 139
[112] OSCE Representative on Freedom of the Media, 'Amsterdam Recommendations on the Freedom of the Media and the Internet' from 14 June 2003 available at https://www.osce.org/fom/13854?download=true accessed on 26 September 2018
[113] *Féret* (n 38) dissenting opinion of judge Sajó joined by judges Zagrebelsky and Tsotsoria
[114] Bibi van Ginkel (n 94)
[115] *Féret* (n 38) dissenting opinion of judge Sajó joined by judges Zagrebelsky and Tsotsoria; John Weckert, 'What is so bad about Internet content regulation?' (2000) 2 Ethics and Information Technology 105, 107-108
[116] *Frisby v Schultz* 487 US 474 (1988); *Consolidated Edison Co v Public Service Comm'n of New York* 447 US 530, 447 US 542 (1980); *Cf Bolger v Youngs Drug Products Corp* 463 US 60 (1983); Raphael Cohen-Almagor, *Speech, Media and Ethics, The Limits of Free Expression: Critical Studies on Freedom of Expression, Freedom of the Press and the Public's Right to Know* (Palgrave 2001) 13
[117] Gargi Chakrabarti and Saahil Dama (n 43)

in mind when assessing the question of imposing an obligation to Internet Intermediaries to delete third-party content online.

The opinion about the impact of the Internet on freedom of expression differs between Europe and US. In Europe, unfortunately, it is believed that Internet may raise a lot of problems for freedom of expression and other human rights and, thus, as being different from traditional media, should be subjected to different measures which would require greater restrictions of freedom of expression.[118] On the other hand, in the US, it is believed that "the interest in encouraging freedom of expression in a democratic society outweighs any theoretical but unproven benefit of censorship."[119]

B. Conclusion

Defining hate speech represents a worldwide challenge, since there is still no universally accepted definition which sets out clear standards, although many have tried to give one. This leads to courts experiencing many problems when adjudicating whether certain expressions indeed amount to hate speech, which could only be solved with proper balancing of all the rights affected. Being similar to offline one, but with certain characteristics resulting from Internet's special features, online hate speech has become a challenge for society. Its boundarylessness, possibility of anonymity and persistence influenced the government to include Internet Intermediaries in its suppression, which represent one of the most important characteristics of online hate speech nowadays. However, the importance of the Internet and its positive impact on the freedom of

[118] Oreste Pollicino and Marco Bassini, 'Free speech, defamation and the limits to freedom of expression in the EU: a comparative analysis' in A. Savin and J. Trzaskowski, *Research Handbook on EU Internet Law* (eds) (2014), p 530
[119] *Reno v ACLU* 521 US 844 (1997), 885

expression cannot be set aside, since it offers a place where people can exchange ideas and engage into debates.

CHAPTER II INTERNET INTERMEDIARIES

A. Introduction

This Chapter will start by defining Internet Intermediaries and listing different types with focusing on the ones who are merely hosting third-party content without engaging into its modification. It will continue with elaborating on different models of Internet Intermediaries liability, namely broad immunity model, safe harbor regime and strict liability regime. From that point it will turn to strict liability model by comparing how is this model implemented in the ECtHR and Germany, and how India shifted away from it by adopting broad immunity model securing freedom of expression protection.

1. Defining Internet Intermediaries

1.1. Defining Internet Intermediaries: General overview

Generally speaking, Intermediary can be defined as third parties who "enable communication of information between two or more different parties".[120] However, when it comes to defining Internet Intermediaries, OECD gives a clear definition whereby it states that Internet Intermediaries represent those entities that "bring together or facilitate transactions between third parties on the Internet by giving access to, hosting, transmitting and indexing content, products and services originated by third parties on the Internet or provide Internet-based services to third parties."[121]

[120] Thomas F Cotter, 'Some Observations on the Law and Economics of Intermediaries' (2006) 1 Michigan State Law Review 67, 68

[121] OECD, 'The Economic and Social Role of Internet Intermediaries' (2010) p 9 <https://www.oecd.org/internet/ieconomy/44949023.pdf> accessed 7 February 2018

A wide range of entities falls under the Internet Intermediaries' definition. Many organizations are trying to provide a list of different types of Internet Intermediaries based on their role.[122] Offered typologies mainly include: *Internet access providers* which allow access to Internet such as e.g. Comcast or free.fr or mobile operators such as Vodafone, T-Mobile, etc., *web hosting providers* which provide web server space for setting up websites, *search engines* such as Google or Yahoo!, *social media platforms*[123] such as Facebook, Twitter, etc.

For the purposes of this thesis, the emphasis will be put only on Internet Intermediaries that host third-party content without modifying it. The fact that some of the Internet Intermediaries, except from hosting the third-party content, may also produce their own, will not influence the thesis focus, as long as those Intermediaries remain to perform the role of the hosting providers. In that manner, the thesis will not focus on pure "content producers"[124], such as news portals which disseminate their own content without offering the opportunity for posting comments from third-parties, but on "active Internet Intermediary"[125], such as news portals which, except from producing their own content, offer the possibility for others to post comments, without being involved in their modification and "passive Internet Intermediary", which only host third-party content without making its own. Thus, to put it simply, the thesis will focus on those *Internet Intermediaries that, irrelevantly to their active role in making their own content or offering other services, host third-party content*, such as social media platforms, news portals, blog owners, and similar.[126]

[122] ibid, p 9; Article 19 (n 6), p 6; CDT, 'Shielding the Messengers: Protecting Platforms for Expression and Innovation' (2012) p 3 <https://www.cdt.org/files/pdfs/CDT-Intermediary-Liability-2012.pdf> accessed 7 February 2018

[123] These platforms are regarded as web 2.0 applications.

[124] UNESCO, 'Fostering Freedom Online: The role of Internet Intermediaries' (2014) p 19 <http://unesdoc.unesco.org/images/0023/002311/231162e.pdf> accessed 7 February 2018

[125] *Delfi AS* (n 10) dissenting opinion of judges Sajo and Tsotsoria, para 11

[126] Article 19 (n 6), p 6

1.2. Defining Internet Intermediaries: ECtHR, Germany and India

Since ECtHR, Germany and India are jurisdictions this thesis is focusing on, this part will elaborate more on the definition of Internet Intermediaries in these jurisdictions. In this manner, the previously mentioned differentiation between "passive Internet Intermediaries" as mere hosting providers and "active Internet Intermediaries" as providers that both host and produce content seems important, due to the definition of types of Internet Intermediaries and their liability regimes given by the E-Commerce Directive, which the ECtHR accepts in its case-law.

EU has made an attempt to set rules for Internet Intermediaries in general by adopting the E-Commerce Directive.[127] Its aim was to enable creation of internal e-market,[128] since before adopting this directive, legislations offered solutions which differed between themselves, thus, making it necessary to create a common rule which would apply equally in all EU countries.[129] However, the E-Commerce Directive has many shortcomings. By trying to regulate e-market and the role of Internet Intermediaries in general, it failed to include all of them, which lead to not applying E-Commerce Directive in some cases, where it should have been applied, e.g. case of *Delfi AS* which will be further elaborated.

Namely, the E-Commerce Directive makes difference among three types of Internet Intermediaries – "mere conduit"[130], "catching"[131] and "hosting"[132] providers and prescribes their liability regime. These Internet Intermediaries provide information society services, which are defined as "any

[127] Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects on information society services, in particular electronic commerce, in the Internal Market, OJ L 178, 17.7.2000
[128] Article 1 of the E-Commerce Directive states that: "This Directive seeks to contribute to the proper functioning of the internal market by ensuring the free movement of information society services between the Member States."
[129] European Commission, 'Liability of Internet Intermediaries in Legal analysis of a Single Market for the Information Society (SMART 2007/0037)' (2011) <https://ec.europa.eu/digital-single-market/news/legal-analysis-single-market-information-society-smart-20070037> accessed 12 November 2018
[130] Directive 2000/31/EC (n 127), Art 12
[131] ibid, Art 13
[132] ibid, Art 14

service normally provided for remuneration, at a distance, by means of electronic equipment for the processing (including digital compression) and storage of data, and at the individual request of a recipient of a service." Such definition of services provided by Internet Intermediaries raises many problems in practice. First, this means that the E-Commerce Directive will apply only to those Internet Intermediaries that provide services for remuneration, including both remunerations gained by users and through commercials. This raises a question whether some Internet Intermediaries which do not profit by users' remunerations or commercial would benefit from liability regime and safe harbor protection offered by Directive.

Moreover, the E-Commerce Directive fails to take into account web 2.0 services, such as Facebook, Twitter and other similar social media networks. These Intermediaries offer a developed package of different services which are not purely passive, nor purely active to fit into one of the categories set by E-Commerce Directive. As not offering only passive services, web 2.0 services may not be regarded as hosting providers, because according to the E-Commerce Directive, hosting providers' activity "consists of the storage of information provided by a recipient of the service". Since web 2.0 services do not only store the information, they do not fall under hosting providers and may not enjoy protection under Article 14 of the E-Commerce Directive which prescribes safe harbor regime – a regime where an Internet Intermediary is exempted from liability if it does not have actual knowledge of illegally of third-party content it hosted and if upon obtaining such knowledge it acts expeditiously to remove the illegal content. This problem is a result of the fact that the directive was drafted at the time when many of the internet services had not yet been developed. Nowadays, the line between transmission services and purely content services has become blurry, and society is faced with question to what extent are some of the Internet Intermediaries neutral.

The way E-Commerce Directive prescribes types of Intermediaries and their liability regimes is important for understanding the ECtHR stance in the judgement which will be elaborated in the following part of the thesis. By not understanding what an internet intermediary is, the ECtHR faced the situation in which it did not exempt from liability online news portal, as it regarded the portal as content provider which did not enjoy safe harbor regime, disregarding its purely passive role in hosting third-party content. As a result, the ECtHR qualified the internet news portal as the editor of the third-party content.

Understanding the nature of Internet Intermediaries is extremely important, especially when setting rules regarding their liability. Since general term of Internet Intermediaries may include many different types of internet service providers, legislators must be careful. In that manner, new law enacted in Germany with an aim to regulate big social networks' liability, does not give a clear definition on Internet Intermediaries whose behavior they want to regulate. The object of the regulation of this act are *"telemedia service providers which, for profit-making purposes, operate internet platforms which are designed to enable users to share any content with other users or to make such content available to the public"*.[133] As it will be shown, the definition is too broad and, as such, may cause many problems in practice.

Regarding India, as one of the three jurisdictions which are the object of this thesis' analysis, the role of Internet Intermediaries is prescribed by the IT Act. The IT Act defines Intermediaries "with respect to any particular electronic records as any person who on behalf of another person receives, stores or transmits that record or provides any service with respect to that record and includes telecom service providers, network service providers, internet service providers, web-hosting

[133] NetzDG Art 1 Sec 1(1)

service providers, search engines, online payment sites, online-auction sites, online-market places and cyber cafes".[134]

When enacting laws which prescribe Internet Intermediaries liability for third-party content they host, it is of extreme importance to carefully define the type of the Intermediary it refers to, assess its nature and provide rules which are sufficiently clear to avoid problems.

2. Models of Internet Intermediaries' liability

As the thesis emphasize is put on the Internet Intermediaries that host third-party content, the liability regimes set forward are as well primarily focused on those Intermediaries. In the literature and in the practice, couple of main models on Internet Intermediaries' liability can be found. The level of Internet Intermediaries' liability for third-party content will depend on the solution adopted on national level or supranational as in case of the EU. Comparison among different jurisdictions reveals three main types of Internet Intermediaries' liability for third-party content they host, including the content which is the subject of this thesis – hate speech. These three types of Internet Intermediaries' liability model are broad/absolute immunity model regime, safe harbor regime and strict liability model regime.[135]

2.1. Broad/absolute immunity regime

Broad immunity model exists in cases Internet Intermediaries are seen as "messengers" who enjoy broad immunity for third-party content they host, without any obligation to monitor such content.[136] Some authors claim that this type of model actually completely shields Internet

[134] Information Technology Act (n 2), Sec 2(1)(w)
[135] Article 19 (n 6)
[136] ibid

Intermediaries from any liability, not requiring them to interfere with any type of illegal content.[137] Anyhow, this model is very rare, while safe harbor regime seems more present.

Some authors mention as a separate model the so called "notice-and-judicial-takedown".[138] It seems more appropriate to fit this model into broad immunity regime, since this model does not require from Internet Intermediaries to monitor the content they host, nor to interfere with any type of illegal content, except in the cases when there is a court order which requires from the Intermediary a certain action. In this situation, Intermediary may be found liable only if it disobeys a court order. India, as a jurisdiction analyzed in this thesis, prescribes a liability regime which fits into broad immunity regime model.

2.2. Safe harbor regime

Safe harbor regime represents most common model of Internet Intermediaries' liability. In line with the safe harbor regime, an Internet Intermediary will be exempted from liability for third-party content if it complies with certain requirements.[139] Opposite to strict liability regime where, as it will be shown, Internet Intermediaries are required to monitor the content they host, in the case of safe harbor regime this obligation does not exist. Instead, Internet Intermediaries are required to act upon notification, which would bring to their knowledge that certain content they hosted was illegal (in the light of the present thesis represented hate speech) and would require from them to act in a suitable manner. What type of notification will be necessary and how should they act to avoid liability, depends on the type of model adopted within this regime. In that manner,

[137] Christina Angelopoulos and Stijn Smet (n 14), 287
[138] ibid, 299-300
[139] Article 19 (n 6), p 7

there are different modifications to this regime, depending on the different legal solutions which can be found, such as notice-and-takedown, notice-and-notice or notice-wait-and-takedown.[140]

Notice-and-takedown system is the most common model and it requires from Internet Intermediaries to take down the allegedly harmful content upon receiving a notification from a user stating the unlawfulness of the content. For many, this model is seen as a perfect balancing of freedom of expression and protection of rights of other and prevention of crime. Namely, the discussion is enhanced online, however, once a certain comment is seen as harmful by users, it may be taken down by intermediary after the user notifies it. The most typical example of notice-and-takedown system implemented through the safe harbor regime can be found in the EU law.

According to the E-Commerce Directive, Internet Intermediaries can be held liable for third-party content which they host, if they fail to fulfil certain requirements. Namely, as already mentioned briefly in the first part of this Chapter, E-Commerce Directive makes difference between "mere conduit", "catching" and "hosting" Internet Intermediaries. Since this thesis deals with Internet Intermediaries that host third-party content, the safe harbor analysis focuses on E-Commerce Directive rules regarding this type of Intermediaries. Internet intermediaries which are regarded as hosting providers according to the Directive, will not be held liable for unlawful third-party content if two conditions are met; first, they do not have actual knowledge about unlawfulness of the content, and second, if they upon obtaining such knowledge act expeditiously to remove such content.[141] In order not to be held liable these Intermediaries should remain passive, meaning that

[140] Christina Angelopoulos and Stijn Smet (n 14)
[141] Directive 2000/31/EC (n 127), Art 14

"their conduct is merely technical, automatic and passive, pointing to a lack of knowledge or control of the data they store."[142]

Additional model through which safe harbor regime is implemented is *notice-and-notice system*. In order to avoid liability, this system requires from Internet Intermediaries that host the third-party content to – upon receiving the notification from the user about alleged unlawfulness of the content – forward the complaint to the user whose content is the subject of complaint or to just notify the complainant if forwarding is not possible, since Intermediaries must not be required to identify users.[143] Applying this system would mean that Internet Intermediaries could be held liable only in case if they do not forward the content removal request from an alleged victim to alleged perpetrator, so the parties could solve the issue directly.[144] This model is a result of the understanding Internet Intermediaries as facilitators between the two parties – the alleged victim and the alleged perpetrator. In case the alleged perpetrator's identity is anonymous and cannot be established or the parties do not manage to solve the dispute, the victim could then seek the judicial help.[145]

At the end, it is worth also mentioning *notice-wait-and-takedown* model, according to which Internet Intermediaries are required to notify the user who posted allegedly unlawful content and wait for response upon receiving a complaint from the complainant.[146] If the user does not respond,

[142] Joint cases *Google France SARL and Google Inc. v Louis Vuitton Malletier SA* C-236/08, *Google France SARL v Viaticum SA and Luteciel SARL* C-237/08 and *Google France SARL v Centre national de recherche en relations humaines (CNRRH) SARL and Others* C-238/08, para 114

[143] Manila Principles on Intermediary Liability: Best Practices Guidelines for Limiting Intermediary Liability for Content to Promote Freedom of Expression and Innovation (2015) prin 3(d) <https://www.manilaprinciples.org> accessed 10 October 2018

[144] Divij Joshi, 'Indian Intermediary Liability Regime Compliance with the Manila Principles on Intermediary Liability' p 17-18 <https://cis-india.org/internet-governance/files/indian-intermediary-liability-regime> accessed 21 October 2018

[145] ibid, p 17-18

[146] Christina Angelopoulos and Stijn Smet (n 14), 296-297

then the Intermediaries may proceed with takedown. The solution is based on the idea that if the user does not respond to the allegation, Internet Intermediaries would be more secure in the decision to remove the content, since in the case that the content was not unlawful, the user would probably object.[147]

Generally, safe harbor regime seems like the best solution for addressing Internet Intermediaries' liability. The importance of safe harbor regime comes from the fact that this model minimizes censorship by private entities and prevents monitoring of the content which may have chilling effect on third parties.[148]

2.3. Strict liability regime

Strict liability regime requires from Internet Intermediaries to constantly and effectively monitor content they host in order to remove it as quickly as possible if it is likely to be unlawful.[149] In other words, Internet Intermediaries will be held liable for third-party content if they do not automatically takedown the unlawful content and prevent harm.[150] Liability can result in facing enormous fines, criminal sanctions, or even withdrawal of business license. Strict liability regime is primarily present in China[151], but with recent developments, ECtHR and Germany introduced the same system within their jurisdictions.

[147] ibid, 296-297
[148] Index on Censorship, 'Index supports referral request in Delfi v. Estonia' (*Xindex*, 14 January 2014) <http://www.indexoncensorship.org/2014/01/index-supports-referral-request-delfi-v-estonia/> accessed 3 October 2018
[149] UNESCO (n 124), p 40
[150] Christina Angelopoulos and Stijn Smet (n 14), 288-289
[151] Qian Tao, 'The Knowledge Standard for the Internet Intermediary Liability in China' (2012) 20(1) International Journal of Law and Information Technology 1

This type of regime is seen as controversial and monitoring as an obligation incorporated in this regime contravenes international standards on liability of intermediaries[152], alters their functionality[153], and violates their freedom of expression.[154] With the recent changes in some of the mentioned jurisdiction, the thesis will, from this point, focus on strict liability regimes in the cases of hate speech in these jurisdictions and the implications of their adopted solutions.

3. Liability solutions in the cases of hate speech adopted in ECtHR, Germany and India

As mentioned, in the past three years many changes took place around the world with respect to regulation of Internet Intermediaries' liability for third-party content. Feeling under pressure to provide security and respect for human rights to their citizenry, driven by the fear imposed by Internet limitless and emphasizing the severity of hate speech, new liability solutions were introduced in Europe. Namely, in 2015 ECtHR rendered a controversial decision and in 2017 Germany enacted a law, both introducing a strict liability regime for Internet Intermediaries in case of online hate speech. On the other hand, while Europe started to be more restrictive, country from other part of the world, India, offered Internet Intermediaries a broad immunity for third-party content.

3.1. ECtHR: *Delfi AS v Estonia* case-study

The most important judgement delivered by the ECtHR which tackled the question of Internet Intermediaries' liability for third-party hate speech content was *Delfi AS v Estonia*[155]. This

[152] Council Directive 2000/31/EC (n 127), Art 15; Council of Europe Declaration on freedom of communication on the Internet (adopted 28 May 2003) prin 6; Case 8611/12, Corte d'Appello di Milano (21 December 2011), Sezione Prima Penale (Italy); Manila Principles on Intermediary Liability (n 143), prin 1(d); Council of Europe Committee of Ministers, 'Appendix to Recommendation CM/Rec(2018)2 Guidelines for States on actions to be taken vis-à-vis internet intermediaries with due regard to their roles and responsibilities' (2018) p 1.3.5.
[153] Case 8611/12 (n 152)
[154] Case C-70/10 *Scarlet Extended SA v Société belge des auteurs, compositeurs et éditeurs SCRL (SABAM)* (ECJ, 24 November 2011), para 50; Case C-360/10 *SABAM v Netlog NV* (ECJ, 16 February 2012), para 51
[155] *Delfi AS* (n 10)

judgement, as already mentioned, introduced the strict liability regime within Council of Europe and sent message to all the Contracting States that in the cases of online hate speech "active"[156] Internet Intermediaries should be regarded as liable for third-party content they merely host if they do not expeditiously delete allegedly hateful content. The case was decided in the last instance by Grand Chamber in 2015 and sparked a lot of controversies since the judgement might have further negative implications on freedom of expression.

Except from ruling about Internet Intermediary liability regime for third-party content, the case is also important with regards to the way in which the Court re-qualified the contested content from defamatory speech to hate speech, widening the scope of the speech which may be regarded as hateful.

3.1.1. Facts of the case

Delfi AS v Estonia concerned the largest Internet news portal in Estonia (hereinafter: the Applicant), which at the end of every article offered a possibility for third-parties to post comments and debate upon different matters contained in the news articles. Delfi AS did not anyhow moderate or edit posted comments; the comments were automatically uploaded online by the users. There was a disclaimer which stated that the users were accountable for the uploaded comments and that any comment which consisted of, *inter alia,* obscene words, threats or incited hostility or violence was prohibited. The Applicant has installed a system of notice-and-take-down, which enabled users to report allegedly harmful content (such as insulting messages and hate speech). Upon the users' notification, the Applicant would remove such content expeditiously.[157]

[156] In the case, the term "active" Internet Intermediary was used to define those Internet Intermediaries that not only host third-party content, but as well provide content of their own. On the opposite, the term "passive" Internet Intermediary was used to define those Internet Intermediaries whose services only consist of hosting the third-party content.
[157] *Delfi AS* (n 10), para 12-14

Additionally, the Applicant installed a system of automatic deletion of comments, which made automatic deletion of obscene words possible.[158] Lastly, there was a possibility for users considering themselves as victims of defamatory or hateful comments to directly notify the Applicant and request their removal. The Applicant, upon such request, would immediately delete contested comments.

In the case at hand, the Applicant published an online article criticizing the public ferry transport service for destroying public ice roads over the frozen sea. This article attracted many third-party offensive comments[159] including personal threats which were directed towards one of the supervisory board members of the public ferry transport service company (hereinafter: Mr. L).[160] As a result, the lawyers of Mr. L requested from the Applicant to remove offensive comments and pay compensation for non-pecuniary damages.[161] The Applicant, on the same day when it received the complaint, deleted all disputed comments, however, refusing the claim for damages.[162] As a result, Mr. L's lawyers filed a civil complaint seeking compensation before national authorities.

In the first instance, the County Court ruled that the safe harbor regime under Information Society Services Act[163] should apply in the case of the Applicant since its role was passive with regards to third-party comments. Namely, the County Court emphasized the fact that Delfi AS was not a publisher of the comments, nor did it have any obligation to edit or monitor them.[164]

[158] ibid, para 13
[159] Some of the comments were later on regarded as hate speech by Grand Chambers decision, as it will be elaborated in the following parts of the thesis.
[160] *Delfi AS* (n 10), para 16-17
[161] ibid, para 18
[162] ibid, para 19-20
[163] Information Society Services Act transposed the European E-Commerce Directive 2000/31/EC in Estonian legal system. As already elaborated in previous subchapter, according to the E-Directive Internet Intermediaries enjoy safe harbor regime when 'their activity is of mere technical, automatic and passive nature' and they do not have 'neither knowledge nor control over the content'. See Recital 42 of the European E-Commerce Directive 2000/31/EC.
[164] *Delfi AS* (n 10), para 23

Mr. L lodged an appeal before Tallinn Court of Appeal, which was later on allowed. The Court of Appeal found that first instance court had erred when it had applied Information Society Services Act and excluded the Applicant's liability for third-party content. Consequently, it quashed the first instance judgement and referred the case back to the County Court.[165]

In renewed proceedings before first instance Court, it was found that the Applicant, as a content provider rather than a passive Internet Intermediary, should be regarded as publisher under Obligation Act and should not enjoy protection under Information Society Services Act, since it does not fall within its scope due to the nature of activities it performs. The County Court found that systems in place, such automatic deletion of obscene comments and notice-and-take-down system, were not enough to protect personality rights of others, thus ruling in favor of Mr. L.[166] This judgement was upheld by Tallinn Court of Appeal. Following, Supreme Court dismissed the Applicant's appeal and upheld the second instance ruling in substance modifying the reasoning in certain parts.[167] In its judgement, Supreme Court specifically elaborated upon the nature of Delfi AS regarding it as "content provider" and emphasizing the economic interest it had in the posting of comments. Supreme Court pointed out that the Applicant (in national proceeding defendant) governed the information in comments by enacting the rules which apply to comments section and by setting systems of control. Thus, it could not be regarded as passive Intermediary which would enjoy protection under Information Society Services Act, but as publisher, which by not preventing the publication of the unlawful content breached the obligation not to cause harm.[168]

[165] Ibid., para 24
[166] ibid, para 26-27
[167] ibid, para 28-31
[168] ibid, para 31

Following the national judgments, the Applicant filed a complaint to the European Court of Human Rights claiming violation of its freedom to impart information guaranteed under the Article 10 of the European Convention.

3.1.2. Chamber judgement

Before the Chamber, there was no dispute about whether there was an interference with the Applicant's right to freedom of expression. However, the Estonian Government claimed that the Applicant should be regarded as publisher of the allegedly defamatory content and that the interference with the Applicants right to freedom of expression was prescribed by law, pursued the legitimate aim of the reputation and rights of others and was justified and proportional, which was contested by the Applicant.[169]

Being satisfied with the applied provision of civil law which prescribed liability of media publishers for defamatory content in their publication, the Chamber did not want to analyze national courts' decision not to apply E-Commerce Directive in the case of Applicant.[170] It concluded that the interference was prescribed by law and pursued a legitimate aim of the protection of the reputation and rights of others. As regards the necessity and proportionality of the interference, the Chamber examined the context of the allegedly defamatory comments, the measures that the Applicant took to prevent and remove these comments, the liability of the third-party users who wrote the comments, and the consequences the Applicant had to bear as a result of domestic proceedings.[171] In conclusion, the Chamber pointed out that the Applicant could have foreseen that the online news article would attract negative comments due to the fact that the article relates to the issues of public interest, but it failed to "exercise a degree of caution to avoid being

[169] ibid, para 61
[170] ibid, para 62
[171] ibid, para 64-65

held liable for damaging the reputation of others".[172] It concluded that available measures put in place by the Applicant were not sufficient and that the Applicant should have *prevented* defamatory content to become public in the first place.[173] In addition the Chamber pointed out that since the Applicant allowed the possibility for third-parties to post comments on its webpage, its liability for those comments represent a natural consequence, especially since it had commercial benefits resulting from comments.[174]

Chamber reached an interesting decision whereby it found no violation of Article 10 although the Applicant deleted the offensive comments as soon as it was notified by Mr. L. At the same time, Chamber declared the online news article as not defamatory nor amounting to hate speech. Seen as controversial, the case was later on referred to the Grand Chamber by the Applicant's request. The referral was supported by 69 international human rights and media organizations, internet companies and academic institutions[175], who expressed their concern about the possible adverse impact of the Chamber judgement on freedom of expression.[176] In the joint letter[177] sent to the ECtHR, the supporters expressed their concern about the lack of any guidance what is expected from Internet Intermediaries such as news portals to do to avoid the liability for third-party content. As well, they pointed out that the judgement may result in censorship and influence Internet Intermediaries in such a way which would make them unwilling to post any sort of content related to questions of public concerns that may attract negative comments. Additionally, out of caution, they might disable the possibility to add comments and by doing so, decrease the public debate.

[172] ibid, para 65
[173] ibid, para 65
[174] ibid, para 65
[175] Some of the companies and organizations were Google, Thomson Reuters, the New York Times, European Newspaper Publishers' Association, Index on Censorship, Greenpeace, the Center for Democracy and Technology, Article 19, etc.
[176] Index on Censorship (n 148)
[177] ibid

Furthermore, the supporters of the joint letter emphasized that the judgement went against CoE standards and EU law, according to which Internet Intermediaries may be held liable only if they do not expeditiously remove harmful content after obtaining knowledge about its illegality. At the end, they also emphasized that the Chamber failed to assess the importance of the E-Commerce Directive and its applicability in the case at hand, since for these authors, it was quite expected that the Applicant would have thought that safe harbor regime applied in its case.

Apart from the mentioned, the Chamber decision as well triggered the negative reaction due to its decision to consent with the application of the rules for publisher in offline world to news portal in online world. This solution cannot be regarded as the one which would respond to challenges of new era of internet technology.[178]

Since the Chamber judgement threatened to introduce strict liability regime for Internet Intermediaries and endanger freedom of expression, as a core right in a democratic society, Grand Chamber decided to hear the case.

3.1.3. Grand Chamber judgement

Before the Grand Chamber, the Applicant claimed that it should not be regarded as publisher, but as Internet Intermediary, thus emphasizing that the tort law could not be applied to the area of new technologies, as Estonian national courts did in domestic proceedings.[179] The law that should have been applied – E-Commerce Directive – prohibited the imposition of liability to Internet Intermediaries for third-party content, and applicant's conduct was fully in line with it.[180] For these reasons, the Applicant claimed that its conviction was not prescribed by law. Additionally, the

[178] ibid
[179] *Delfi AS* (n 10), para 69-70
[180] ibid, para 69

Applicant claimed that the conviction was not necessary in a democratic society since, *inter alia*, the article that attracted contested comments was neutral, the Applicant deleted the comments as soon as it found out about them, actual authors of the comments could have been tracked and there was a European consensus that Internet Intermediaries should not be held liable for third-party content.[181]

On the other hand, Government once again emphasized that Delfi's conviction based in tort law was lawful since the offensive comments were integral part of the news articles which could only be modified and deleted by the Applicant, and not by the users. For this reason, Delfi AS could not have been regarded as caching nor hosting service provider, thus making E-Commerce Directive non-applicable in the case at hand.[182] As regards necessity in a democratic society, the Estonian government pointed out that the Applicant was a discloser of anonymous comments on its news articles and that it carried a certain level of liability by allowing anonymous third-party comments which only could have been deleted or modified by the Applicant.[183] It added, as well, the fact that anonymous authors could not have been revealed.[184]

The Helsinki Foundation for Human Rights as a third-party intervener emphasized that Delfi AS is both content and host provider. It provided content in the form of news, at the same time hosting third-party comments. Sharing the same view, judges Sajo and Tsotsoria in their dissenting opinions regarded the Applicant as "active Internet Intermediary"[185]. Additionally, The Helsinki Foundation for Human Rights pointed out that the power to moderate user-generated content, in

[181] ibid., para 72-78
[182] ibid, para 84-85
[183] ibid, para 89-93
[184] ibid, para 91
[185] Although regarded by Chamber and Grand Chamber as active content provider which do not represent hosting Intermediaries and does not enjoy exemption from liability under E-Directive, in their dissenting opinion of judges Sajo and Tsotsoria use the term active Internet Intermediary to describe this type of Intermediaries, that contrary to the passive ones, not only host third-party content, but also post their own.

sense of deleting it or preventing its publication, cannot be regarded as having editorial control, because online news portals cannot be treated like traditional media.[186]

In its judgment, the ECtHR recognized the fact that Delfi AS is, due to its nature, somehow different from the publisher of the printed media publication who is editor of all content appearing in his publication, but at the same time, the Court emphasized the fact that the Applicant had economic interest in the publication of the comments in the same way as publishers of traditional media have.[187] By providing, for economic purposes, a platform for user-generated comments on previously published content where third-parties engage in spreading "clearly unlawful speech", the Court made a differentiation between Delfi AS and other Internet Intermediaries that would otherwise enjoy protection from being held liable for third-party content.

Following the three-tier test, the Grand Chamber while deciding about *lawfulness of the interference* applied the principle of subsidiarity not wanting to enter into discussion about the law which should have been applied by the domestic courts. Instead, it looked whether the Applicant could have foreseen the possibility of being held liable. Thus, the Grand Chamber concluded that the interference was prescribed by law and that the Applicant as media published could have foreseen that it could have been held liable for uploading clearly unlawful comments.[188] Such ruling once again clearly expressed the Grand Chamber's opinion that the Applicant should have monitored all third-party content. As regards *legitimate aim* of protecting the reputation and rights of other, there was no dispute between the parties. However, as regards the third part of the test applied by the ECtHR, the Grand Chamber ruled that the interference with the Applicant's freedom

[186] *Delfi AS* (n 10), para 94-95
[187] ibid, para 112
[188] ibid, para 128

of expression was *necessary in a democratic society*. The Grand Chamber assessed the same aspects as the Chamber did in its judgement.

The Grand Chamber went on by ruling that although the news article the Applicant posted was balanced, the third-party comments were not, and the Applicant could not have been regarded as purely passive Internet Intermediary since it was involved in making those comments public.[189] Such conclusion was drawn from the fact that only the Applicant had the possibility to delete third-party comments, while their actual authors could have not. However, it is quite debatable whether this would be enough to consider an Internet Intermediary as actively involved in making content public.

While assessing the necessity of the interference, the Grand Chamber as well took a look at the possibility for authors of the comments to be held liable. It concluded that it was often not possible to reveal the identity of the real authors of the comments, and since the Applicant did not install any measure which would enable potential victims to identify them and bring a claim, the interference was necessary.[190] It seems the Court is suggesting that, *inter alia,* the real name policy would have protected the Applicant for being held liable, which represents a controversial solution, as it will be elaborated in the following chapter of the thesis.

As regards measures taken by the Applicant to remove the hate speech comments, the Grand Chamber concluded that the Applicant had sufficient degree of control over these comments. The Grand Chamber recognized installed notice-and-take-down system and system of automatic deletion of obscene words, as well as self-initial removal of inappropriate comments by administrators. However, it pointed out that these were insufficient, primarily the filtering

[189] ibid, para 144-146
[190] ibid, para 150-151

mechanisms which failed to detect hate speech comments. In addition, it emphasized Delfi's duty to care and ruled that, as a large news portal run on a commercial basis, the Applicant should have employed more efficient mechanism, since it could have not been expected from the victim to monitor content. Thus, it concluded that the Applicant failed to delete the content without delay after their publication.[191] Such ruling completely disregards the fact that the applicant did delete all the comments after it was notified by the injured party which was six weeks after their publication. As well, it raises many issues with regards to hate speech and the question how a filter mechanism could possibly make difference between legitimate speech and hate speech, especially taking into consideration the comments written on Delfi's news article, which are not considered as hate speech by many authors. At the end, it opens a question what mechanism would be satisfactory for the Court, since the Applicant already had many installed.

At the end, the Grand Chamber agreed with Chamber and concluded that the interference was *proportionate*, since the amount of damages awarded was not high and the Applicant remained to be one of the largest news portals in Estonia. Additionally, it continued working normally without being forced to change its business model.

The Grand Chambers message was clear: in the cases of hate speech an Internet Intermediary such as news portal has to take expeditious measures to remove allegedly harmful content even without notice from the alleged victim or any third party![192] Thus, by fifteen votes to two it reached a decision in which it found no violation of Article 10 of the Convention and that Internet Intermediaries such as news portals should monitor and delete hate speech content online.

[191] ibid, para 153
[192] ibid, para 159

It is important to add one more issue regarding the qualification of contested comments. The Grand Chamber made a step further than Chamber, by ruling that the impugned comments were "hate speech and speech that directly advocated acts of violence".[193] This leads back to the discussion about the definition of hate speech and what does it actually entail. As already mentioned in the first Chapter, a line between freedom of expression and hate speech is very hard to establish. It seems hard to believe that comments like "bastards!!! Ofelia also has an ice class, so this is no excuse why Ola was required!!!"[194] or "…a question arises whose pockets and mouths he has filled up with money so that he's acting like a pig from year to year…"[195] written under the article that raises public issues may be regarded as hate speech. Even Chamber which is consisted of eminent judges, in its decision did not qualify it as hate speech.

Although the Grand Chamber limits the influence of the judgment only to "professionally managed internet news portal run on a commercial basis" excluding social media networks and other fora used for "hobby", it seems still as something that would lead to "collateral damage of freedom of expression online". [196]

In the joint concurring opinion, judges Raimondi, Karakas, De Gaetano and Kjolbro pointed out that there were two possible readings of the Estonian Supreme Court judgement as regards Delfi's liability for not preventing the unlawful comments from being published. The concurring judges emphasized that the Grand Chamber understood that the Applicant's liability arises from not removing the unlawful comments subsequently and expeditiously after their publication, and not

[193] ibid, para 117
[194] ibid, para 18(16)
[195] ibid, para 18(19)
[196] Dirk 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) <https://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 10 October 2018

from not preventing their publication *a priori*.[197] However, it seems that both understanding of liability will lead to the same consequence – the Applicant and other Internet Intermediaries will have to monitor all the content in order to delete it expeditiously enough, which represent an excessive burden as it will be explained in the next chapter.

3.1.4. Possible negative implications of the judgement: Dissenting opinions

At the end, it is important mentioning that two judges dissented and voted for violation of Article 10. In their dissenting opinion they especially emphasized the fact that the Court's role is to safeguard human rights observing the general context and determine the issues on public policy grounds, not only to provide individual relief, which the Court failed to do in the case at hand.[198] Additionally, they emphasized that the Court failed to take into account that in majority High Contracting states and other leading world democracies a safe harbor regime is provided for Internet Intermediaries, primarily through imposing notice-and-take-down model.[199] In that sense, the Court disregarded the concept of "actual knowledge" which is widespread concept, and introduced the strict liability regime based on pure "promptness".

By emphasizing that the ruling will not apply to social media network or other forum on the Internet where *"the content provider may be a private person running the website or a blog as a hobby"* the Court is not actually setting proper safeguards, especially having in mind the rightful point made by dissenters whereby freedom of expression cannot be regarded as a matter of a hobby.[200]

[197] *Delfi AS* (n 10), concurring opinion of judges Raimondi, Karakas, De Gaetano and Kjolbro, para 4-5
[198] ibid, dissenting opinion of judges Sajo and Tsotsoria, para 4-6
[199] ibid, dissenting opinion of judges Sajo and Tsotsoria, para 7
[200] ibid, dissenting opinion of judges Sajo and Tsotsoria, para 9

Additionally, dissenters tackled separately the question of lawfulness of the interference, emphasizing that the Court failed to provide a reasonable justification why did it opt for higher level of liability set down in Estonian Civil Code and not applying the EU directive which had the supremacy as *lex specialis*.[201] They pointed out that comments on the Applicant's news portal were user-generated, as such making the safe harbor regime under the E-Commerce Directive applicable in present case, notwithstanding the commercial nature of the data storage that was used by the Court to claim otherwise. Additionally, discussing about the same issue, the dissenters claimed that this law was not foreseeable even for legal advisers since even Cyprus Court in 2013 in the case of news portal publisher *Papasavvas* had to ask CJEU for preliminary ruling upon applicability of the E-Commerce Directive, and this question was clarified by the CJEU eight years after the case at hand has been brought before Estonian courts. As well, neither Civil Code Act nor the Law of Obligations act as too general could have been regarded as foreseeable, since no legal council could have thought that an online news portal could be liable for third-party comment it was not aware of under the liability rules which applied to editors who were aware of whole publications.[202]

Dissenters pointed out that the ECtHR failed to struck a fair balance, since it did not took into account different factors, such as *inter alia* possibility of adopting less intrusive measure, the fact that the interference concerns journalism which triggers strict scrutiny, that the matter of the article was one of the public interest and that the possibility for commenting by third-parties enhanced public debate concerning public matter and enabled people to receive and impart information, and that the basic principle in ECtHR case-law was that "punishment of journalist who simply

[201] ibid, dissenting opinion of judges Sajo and Tsotsoria, para 17
[202] ibid, dissenting opinion of judges Sajo and Tsotsoria, para 20

disseminates the statements made by others in an interview hampers press contribution to discussion matter of public debate and should not be envisaged unless there are particularly strong reasons for doing so".[203]

Dissenters pointed our that the economic interest is not enough to regard an active Internet Intermediaries as publishers nor the fact that it can exercise certain level of control over the comments, since the differences between publishers and Internet Intermediaries such as Delfi AS are huge. [204] These differences include lack of personal control over the users who post comments without the decision of the Internet Intermediary, so Internet Intermediary does not know the content of an comment in advance before the posting.[205] They emphasized the importance of *Delfi AS* in enhancing public debates related to public matters and pointed out that its protection should not be diminished.[206]

Dissenters as well criticized the reasoning of the Court whereby it found that filtering mechanism, as a "simple" mechanism, failed, not discussing about any possibly less intrusive measure which would satisfy the goal of removing hate speech online except removal without delay, which would mean introduction of pre-monitoring obligation that goes against all international standards.[207]

At the end, they, as well, rightly point out that even the victim did not follow the article and did not see the allegedly harmful content earlier than six weeks after its publication which was directly related to the victim's company. Given the fact that the contribution of the victim is also something that should be considered in standard liability cases, how come it was not included here.

[203] ibid, dissenting opinion of judges Sajo and Tsotsoria, para 39-40
[204] ibid, dissenting opinion of judges Sajo and Tsotsoria, para 28-32
[205] ibid, dissenting opinion of judges Sajo and Tsotsoria, para 31
[206] ibid, dissenting opinion of judges Sajo and Tsotsoria, para 28-30
[207] ibid, dissenting opinion of judges Sajo and Tsotsoria, para 36

To conclude, such ruling whereby an Internet Intermediary is required to introduce measures – other than notice-and-take-down and prior automatic filtering – which would prevent potentially defamatory content to become public represents nothing more that general obligation to monitor content, which goes against international standards, for reasons which will be explained in the following chapter.

3.1.5. After Delfi AS v Estonia

After *Delfi AS*, the case concerning defamatory third-party comments on a self-regulatory non-commercial body MTE and a commercial Internet news portal Index.hu reached the ECtHR. The case dealt with the real estate company who brought a claim against these two Internet Intermediaries for offensive anonymous third-party comment which offended its reputation.[208] Contrary to Delfi judgment, the Court here concluded that by holding Internet Intermediaries liable for third-party defamatory content there was a violation of Article 10. The ECtHR made differentiation between the two cases by pointing out that in Delfi judgement the severe impact of the impugned comments, namely the fact that they constituted hate speech and direct call for violence, justified the interference with the Applicant's right to impart information. On the other hand, insulting comments directed towards business policy of a corporate company did not justify the interference. Additionally, what is interesting enough is the fact that in Delfi judgement, the Applicant removed the comments immediately upon the victim's request, while in the *MTE and Index.hu Zrt v. Hungary* the injured company directly addressed the court without requiring removal of the offensive comments, and Internet Intermediaries removed them after being notified about civil proceedings that were initiated against them.

[208] *Magyar Tartalomszolgáltatók Egyesülete And Index.Hu Zrt v Hungary*, App no 22947/13 (ECtHR, 2 February 2016)

National court did not regard these portals as intermediaries and it applied rules about objective liability for operators of the websites for their publication set in the Hungarian Civil Code. MTE and Index.hu claimed before ECtHR that liability could have been avoided only if they applied pre-moderating system or they completely excluded possibility of commenting, which would be harmful towards freedom of expression.

ECtHR, while discussing the issue of necessity of interference, applied the same test as in Delfi. Namely, it looked into the context of the impugned comments, the liability of the authors of the comments, the measures taken by the Applicant, the consequences of the comments for the injured party and the consequences for the applicants.[209] Interestingly, here the Court called upon the earlier case-law which put emphasize on the importance of the journalistic activity and the principle which, due to contribution that the press has on discussion of matters of public interest, requires existence of "particularly strong reasons" to "punish a journalist for assisting in the dissemination of statements made by another person in an interview".[210] The exercised balancing in this case was in favor of an Internet Intermediary.

It stated that "*such liability may have foreseeable negative consequences on the comment environment of an Internet portal, for example by impelling it to close the commenting space altogether. For the Court, these consequences may have, directly or indirectly, a chilling effect on the freedom of expression on the Internet."*[211] It seems strange that the Court failed to apply this same standard in the *Delfi* judgement. Quite opposite, it put even bigger burden on the Internet Intermediary stating that due to severe nature of the impugned comments, even a notice-and-take-

[209] ibid, para 69
[210] ibid, para 79
[211] ibid, para 86

down system which resulted into expeditious content removal upon user's notification was not enough in that case.

This judgement raises important questions, among them why the Grand Chamber in Delfi judgement requalified the comments from defamatory to the ones inciting hatred and *direct* call for violence, when comments by their nature did not differ from the ones posted in *MTE and Index.hu Zrt* case. It seems unbelievable that the Court in the Delfi decision failed to take into account "*the specificities of the style of communication on certain Internet portals*" and that the comments on the *Delfi* platform in the same manner as the ones in *MTE* judgement, *"although belonging to a low register of style, were common in communication on many Internet portals, which reduces the impact that can be attributed to those expressions".*[212] But even if they did qualify those comments as inciting hatred and direct call for violence, does hate speech indeed requires imposition of monitoring obligation to Internet Intermediaries? The *MTE and Index.hu Zrt* and all the differentiation find between that judgement and *Delfi*, seems more like an effort to justify the solution previously adopted by the Grand Chamber, which faced many critics among scholars and practitioners.

Apart from the case of *MTE and Index.hu Zrt*, one additional case that reached ECtHR which tacked the question of Internet Intermediaries' liability was *Pihl v Sweden*.[213] However, this case did not address the issue of hate speech which was repeated once again in this decision. Hateful nature of comments, or in court's words "clearly unlawful content" clearly represents one of the crucial factors for assessing Internet Intermediaries' level of liability within CoE. The case dealt with the blog post that accused the Applicant to be a member of a Nazi party and an anonymous

[212] ibid, para 77
[213] *Pihl v Sweden*, App no 74742/14 (ECtHR, 7 February 2017)

offensive comment written on it.[214] Mr. Pihl (the Applicant) by commenting on the blog post requested the removal of the whole blog post and offensive comment as well. Immediately, the day after the request[215], the blog operator deleted both the post and the comment and wrote a new post where it admitted that previous post was inaccurate, and the mistake was done to the Applicant. Notwithstanding, the Applicant sued the blog operator for defamation and the fact that the post and the comment were deleted only nine days after publication. However, national court ruled in favor of the blog operator. The Applicant then brought the case before the ECtHR, which found the case manifestly ill-founded, since in its view, domestic courts acted within their margin of appreciation in striking the fair balance between competing rights of the applicant and Internet Intermediary.

However, here again the same questions as in the *MTE* case could be raised. It seems hard to believe that the comment such as "That guy Pihl is also a real hash-junkie according to several people I have spoken to"[216] differs much from the comment appearing in *Delfi:* "bastards!!! Ofelia also has an ice class, so this is no excuse why Ola was required!!!"[217]

All these judgements lead to one conclusion: only if an Internet Intermediary starts pre-monitoring *all* of the third-party content it hosts in order to find the content which constitutes "incitement to hatred and direct call for violence" on its own initiative and expeditiously delete it, it can be exempted from liability.[218] Thus, the *Delfi* judgement and all the following judgements endorsed

[214] The comment was following: "That guy Pihl is also a real hash-junkie according to several people I have spoken to".
[215] This was nine days after the blog post and the comment.
[216] *Pihl* (n 213) para 4
[217] *Delfi AS* (n 10), para 18(16)
[218] Dirk 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) <https://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 11 October 2018

the Internet Intermediaries' strict liability regime in the cases of hate speech and introduction of obligation to monitor within the Council of Europe states.

3.2. Germany: The Act to Improve Enforcement of the Law in Social Networks

3.2.1. Reasons for adopting the Act

Apart from the ECtHR, some states have started introducing laws which directly impose strict liability to Internet Intermediaries in the cases of online hate speech. Facing the problem of fake news and online hate speech, imposing strict liability on Internet Intermediaries came as the easiest solutions for Germany. In that manner, in 2017 Germany introduced an Act to Improve Enforcement of the Law in Social Networks[219] (hereinafter: Network Enforcement Act) which imposed a strict obligation to Internet Intermediaries to take down "manifestly unlawful content" within 24 hours and "all unlawful content" within one week of receiving complaint by any party, if they do not want to face huge fines reaching up to 50 million euros.[220]

The Network Enforcement Act uses term Social Networks, since its primary goal was to regulate the procedures and obligations of social media platforms in cases of online hate speech and fake news complaints. When adopting the named act, Germany emphasized that "verbal radicalization is the first step towards physical violence"[221] and since the Internet Intermediaries, primarily social media platforms, failed to voluntarily delete criminally punishable content, legislative action was

[219] NetzDG
[220] ibid, Art 1 Sec 3 and 4
[221] Answers to the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression in regard to the Act to Improve Enforcement of the Law in Social Networks (Network Enforcement Act), provided by the Federal Government of Germany (2017) <https://www.ohchr.org/Documents/Issues/Opinion/Legislation/GermanyReply9Aug2017.pdf> accessed 2 November 2018

in their opinion necessary.[222] However, it is disputable whether the solution adopted by the legislation is indeed a good one, which will be further elaborated.

Germany wanted to achieve its aim to expeditiously clean the Internet from hate speech by introducing a more satisfying level of content removal through the obligation imposed to Internet Intermediaries to remove content in 24 hours or one-week period and to issue periodical reports about the removal.[223]

3.2.2. Object of the regulation

As already mentioned, the Act imposes obligation to Social Networks, which are defined as *"telemedia service providers which, for profit-making purposes, operate internet platforms which are designed to enable users to share any content with other users or to make such content available to the public"*.[224] Although it specifically excludes platforms that offer journalistic content and platforms which enable individual communication,[225] such broad definition still includes different types of Intermediaries, not only social media networks, as it was its primary objective. Namely, it includes games' providers as well, since they often offer communication tools,[226] even though they do not enhance public debate in the same way as social media networks and similar Internet Intermediaries do, nor are they used for online hate speech dissemination to the same extent as other Intermediaries.[227]

[222] ibid
[223] Gerald Spindler, 'Internet Intermediary Liability Reloaded – The New German Act on Responsibility of Social Networks and its (In-) Compatibility with European Law' (2017) 8 JIPITEC 166, 166-167
[224] NetzDG Art 1 Sec 1(1)
[225] ibid, Art 1 Sec 1(1)
[226] During online video games, players in multiplayers game can interact by using chat opportunities or join video calls.
[227] BIU – Bundesverband Interaktive Unterhaltungssoftware, 'Opinion on the draft bill of the German Federal Ministry of Justice and Consumer Protection regarding an act for improving law enforcement on social networks (Netzwerkdurchsetzungsgesetz, NetzDG)' (2017) <http://ec.europa.eu/growth/tools-databases/tris/en/search/?trisaction=search.detail&year=2017&num=127> accessed 31 October 2018

The Act also prescribes that it will not apply to social networks that have less than two million registered users in Germany.[228] This opens a question whether there should be a line drawn between registered and active users, since many users remain registered although they stop using the platforms. One of the examples is MySpace, which was the largest social media network in the world between 2005 and 2009. However, at the moment, although many users did not delete their accounts, it is barely used. This offers place for different proposals, such as amending the provision in a manner which would as a parameter include only active users.[229] Additionally, the problem with this parameter is the situation when many people have more than one registered account on the same platform or when accounts are fake and run by software, so called bots.[230]

At the end, it is unclear what was meant by "registered users in Germany". The Act does not provide an answer to the question whether the two million registered users in Germany refers to the number of users who were in Germany at the time of registration or users who are present in the Germany at the time of use.[231] Additionally, user may use VPN system to hide its IP address, thus, making it impossible to find out whether the user is registered in Germany.[232]

3.2.3. Unlawful content under the Act and the obligation to remove it under relevant time frame

As mentioned, the Network Enforcement Act calls upon Internet Intermediaries to delete unlawful content, which may be "manifestly unlawful" and "unlawful". According to the Act, unlawful content is the one that violates certain provisions set down in German Criminal Code; it has to fulfil requirement of the named offences under German Criminal Code and it must not be

[228] NetzDG Art 1 Sec 1(2)
[229] BIU – Bundesverband Interaktive Unterhaltungssoftware (n 227)
[230] EDRi, 'Recommendations on the German bill "Improving Law Enforcement on Social Networks" (NetzDG)' (2017) <https://edri.org/files/consultations/tris_netzdg_edricontribution_20170620.pdf> accessed 1 November
[231] BIU – Bundesverband Interaktive Unterhaltungssoftware (n 227)
[232] EDRi (n 230)

justified.[233] However, the list of unlawful content is quite broad and includes many different offences which require different level of protection.[234] It includes as well criminal defamation and insult, which is, among other offences[235], considered problematic, since criminal penalties represent disproportionate means for protecting reputation rights of others.[236] Additionally, it includes "incitement to hatred", which is prescribed by Article 130(1) of the German Criminal Code[237], and which is considered as non-complainant with the ICCPR and the ECHR. Namely, the criminalized behavior does not consist of advocacy of hatred constituting incitement to discrimination, hostility or violence, thus making the criminal punishment disproportionate.[238]

As mentioned, the Network Enforcement Act prescribes the time frame of 24 hours for Internet Intermediary to act and take down "manifestly unlawful content" and one week for taking down "all unlawful content". The deadline starts counting after receiving complaint by any party. The

[233] NetzDG Art 1 Sec 1(3) reads as follows: "Unlawful content shall be content within the meaning of subsection (1) which fulfils the requirements of the offences described in sections 86, 86a, 89a, 91, 100a, 111, 126, 129 to 129b, 130, 131, 140, 166, 184b in connection with 184d, 185 to 187, 241 or 269 of the Criminal Code and which is not justified." Offences from Criminal Code listed down in this Section are following: dissemination of propaganda material of unconstitutional organizations (Section 86), using symbols of unconstitutional organizations (Section 86a), preparation of a serious violent offence endangering the state (Section 89a), encouraging the commission of a serious violent offence endangering the state (Section 91), treasonous forgery (Section 100a), public incitement to crime (Section 111), breach of the public peace by threatening to commit offences (Section 126), forming criminal and terrorist organizations, domestically and abroad (Section 129 to 129b), incitement to hatred (Section 130), dissemination of depictions of violence (Section 131), rewarding and approving of offences (Section 140), defamation of religions, religious and ideological associations (Section 166), distribution, acquisition, and possession of child pornography (Section 184b), distribution of pornographic performances by broadcasting, media services, or telecommunications services (Section 184d), insult and defamation (Section 185 to 187), causing the danger of criminal prosecution by informing on a person (Section 241) and forging of data intended to provide proof (Section 269).

[234] Opinion of David Kaye Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression regarding the German Draft Law Netzdurchführungsgeset <https://www.ohchr.org/Documents/Issues/Opinion/Legislation/OL-DEU-1-2017.pdf> accessed 29 June 2018

[235] For all observations see: Article 19, 'Germany: Draft Bill on the Improvement of Enforcement of Rights in Social Networks' (2017) p 12-14 <https://www.article19.org/data/files/medialibrary/38723/170426-Germany-Hate-Speech-Law-Draft-Analysis.pdf> accessed 29 June 2018,

[236] ibid, p 12

[237] The Article 130(1) reads as follows: „Whosoever in a manner capable of disturbing the public peace... incites hatred against a national, racial, religious group or a group defined by their ethnic origins, against segments of the population or individuals because of their belonging to one of the aforementioned groups or segments of the population or calls for violent or arbitrary measures against them... shall be liable to imprisonment..."

[238] Article 19 (n 235) p 14

24 hours' time limit for removal of manifestly lawful content to be exceeded in the case Internet Intermediary reaches an agreement with the competent law enforcement authority about such extension.[239] As regards one-week period – the time limit set for the removal of the "all unlawful content" – this period may be exceeded if Internet Intermediary, due to the necessity of obtaining more factual proofs decides to give an opportunity to the third-party to respond to the complaint before it reaches a decision.[240] As well, this time limit may be exceeded if Internet Intermediary decides to refer the case to the recognized self-institution, which will then be in charge of deciding about the unlawfulness of the content.[241] This body is seen as independent and competent body consisting of expert analysist.[242] However, this decision for referral is optional. As well, it is worth mentioning that in complex situations when it has to be proven whether certain content is a statement of fact or a value judgment, difficulties may be encountered even by courts which may rich completely different decisions, thus making seven days deadline for Internet Intermediaries and even referral to certain self-regulatory body disputable.[243]

Except from the general time frame set by the Network Enforcement Act, it seems important to emphasize that the Act also requires from Internet Intermediaries to take *immediate* note of the complaint and check whether the content is unlawful.[244] Usage of the term "immediate" may cause a lot of additional problems in practice and undoubtedly poses an undue burden on Internet

[239] NetzDG Art 1 Sec 3(2)(2)
[240] ibid, Art 1 Sec 3(2)(3)(a)
[241] ibid, Art 1 Sec 3(2)(3)(b)
[242] Answers to the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression in regard to the Act to Improve Enforcement of the Law in Social Networks (Network Enforcement Act), provided by the Federal Government of Germany (n 221)
[243] UNESCO, 'Comments on the Draft for an Act improving Law Enforcement on Social Networks (NetzDG)' (2017) <https://www.hans-bredow-institut.de/uploads/media/default/cms/media/4c70991cc1654caa2b47b509bad7bd1328824391.pdf> accessed 31 October 2018
[244] NetzDG Art 1 Sec 3(2)(1)

Intermediaries, who, having as much users as prescribed by the Act, may not be able to respond to their imposed obligations that quickly.

3.2.4. Obligation to report and preliminary court ruling about unlawfulness of the content as possible safeguards

As a result of many critics, Germany tried to prescribe some safeguards which would enable protection of freedom of expression and other relevant rights. In that manner, the Act prescribes Internet Intermediaries' reporting obligation as a way to increase transparency of the whole process of content removal.[245] In other words, Internet Intermediaries who receive more than 100 complaints per year would need to issue reports twice a year, which would include the manner in which they handled unlawful content on their platform.[246] The Act prescribes a detailed list of information which needs to be included in the report.[247] Although being a good idea, some have emphasized that it may happen that most of the removed content could be classified as in breach of "community rules", and not as criminal unlawful content, thus making this obligation irrelevant and not reaching its purpose.[248]

Additionally, as one of the possible safeguards, the Network Enforcement Act prescribes the need for a preliminary court ruling about unlawfulness of the content prior to Internet Intermediary is being fined.[249] Although one more good idea in theory, in practice this would not give significant results. Namely, this does not safeguard the fact that Internet Intermediaries will still delete more content than needed while being threatened by huge fines.

[245] ibid, Art 1 Sec 2
[246] ibid, Art 1 Sec 2(1)
[247] ibid, Art 1 Sec 2(2)
[248] EDRi (n 230)
[249] NetzDG Art 1 Sec 4(5)

3.2.5. Accordance with the E-Commerce Directive

E-Commerce Directive represents a *lex specialis* in comparison to the Network Enforcement Act adopted by the Germany. As such, it should not anyhow contravene it. However, the reality is different, and several segments of the Network Enforcement Act go against the mentioned EU Directive.

First of all, regarding the territorial application, Network Enforcement Act makes the Internet Intermediaries' liability possible even in cases the Intermediaries headquarters are not based in Germany as long as the users are Germans. As well, Section 4(3) prescribes that the offences will be sanctioned even if they are not committed in Germany. Since the Act can be applicable to Internet Intermediaries seated in other EU member states, this solution is seen as contravening Article 3 of the E-Commerce Directive, which prescribes country of origin principle with an aim of harmonizing the legal framework for all Internet Intermediaries in the European Union.[250]

Network Enforcement Act, as well, contravenes Article 14 which prescribes exemptions from Internet Intermediaries liability for third-party content they host in cases they act expeditiously to remove illegal content upon obtaining actual knowledge about it.[251] Firstly, Network Enforcement Act, instead of referring to "expeditious removal", sets extremely short time frame of 24 hours and seven days in which Internet Intermediaries should act in order not to be held liable for third-party content and instead of calculating the beginning of the time period from the moment of obtaining "actual knowledge" about it, it starts calculating from the moment of receiving a complaint.[252] Not using the term expeditious removal can *inter alia* lead to all EU member states introducing completely different time frames and can endanger the harmonization of legal framework of

[250] Gerald Spindler (n 223), 167
[251] EDRi (n 230)
[252] Gerald Spindler (n 223), 171-174

member states. As regards using the moment of receiving the complaint instead of "actual knowledge" can cause many problems, since these terms do not match. Internet Intermediary can receive a complaint at one period of time, but it can obtain an actual knowledge of the content of the complaint later on.[253] As well, "actual knowledge" refers to obtaining a knowledge about unlawful nature of the content, not only about its existence.[254]

Additionally, the Network Enforcement Act contravenes Article 14 of the E-Commerce Directive as it defines the Internet Intermediaries in a different manner. Namely, Article 14 of the E-Commerce Directive does not make difference between small and big Internet Intermediaries.[255] Additionally, the privileges and obligations provided in Article 14 apply to all kind of illegal activities, while Network Enforcement Act deals primarily with hate speech and fake news.

3.2.6. Possible negative implications of the Act

By not providing enough precise definition of Social Network, nor any guidance how a private entity should make a difference between manifestly unlawful, unlawful and lawful content[256], and taking into account many other shortcomings[257], the Act has been heavily criticized as reckless move of Germany endangering freedom of expression as one of the core rights of democratic society. Being afraid to face huge fines, German law created a possibility for Social Networks to delete lawful content out of caution and exercise over-censorship.[258] At the same time it is violating

[253] ibid, 172-173
[254] ibid, 173
[255] ibid, 175
[256] Article 19 (n 235) p 16
[257] See e.g.; OSCE, 'Legal Review of the Draft Law on Better Law Enforcement in Social Networks' (2017) <https://www.osce.org/fom/333541?download=true> accessed 29 June 2018; Opinion of David Kaye Special Rapporteur (n 29); Article 19 (n 235); Emma Llanso, 'German Proposal Threatens Censorship on Wide Array of Online Services' *CDT* (7 April 2017) <https://cdt.org/blog/german-proposal-threatens-censorship-on-wide-array-of-online-services/> accessed 29 June 2018
[258] Emma Llanso (n 257); Article 19 (n 235)

international standards for promoting free speech online.[259] It requires from Internet Intermediaries as private actors to play the role of the judge and decide which content is lawful and which is not, without giving any clear guidelines.[260] Although state has responsibility to protect its nation from unlawful conducts, this does not give the authorization to States to place upon private parties the duty to regulate freedom of expression.[261]

Additionally, the Act has been criticized for failing to address the liability of the authors of unlawful content by putting emphasis only on the liability of Internet Intermediaries, who are bare hosting providers without the possibility to modify the third-party content. Prescribing Internet Intermediaries' liability as a tool for ensuring removal of allegedly unlawful content from the Internet will not anyhow affect the third-party's behavior by deterring them from posting similar content in future.[262]

By introducing this Act, Germany set a model for other countries which saw this as a great opportunity to justify their "non-democratic" laws and limit Internet debate. [263] In that manner, similar laws to German are being introduced in other countries such as, *inter alia*, Russia, where social networks are obliged to act within 24 hours from notification[264] and Kenya, where new

[259] Then-Special Rapporteur for Freedom of Expression, Frank La Rue in its Report from 2011 noted that "*Holding intermediaries liable for the content disseminated or created by their users severely undermines the enjoyment of the right to freedom of opinion and expression, because it leads to self-protective and over-broad private censorship, often without transparency and the due process of the law."* See also Manila Principles on Intermediaries' liability that prescribe that *"Intermediaries must not be required to restrict content unless an order has been issued by an independent and impartial judicial authority that has determined that the material at issue is unlawful."* Additionally, Article 14 of the E-Commerce Directive (2000/31/EC) which Germany transposed provides a liability exception for online intermediaries, when they act expeditiously to remove illegal content, according to the notice-and-take-down procedure.

[260] Article 19 (n 235), p 16

[261] Opinion of David Kaye Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression regarding the German Draft Law Netzdurchführungsgeset, (n 234), p 4

[262] Article 19 (n 235), p 17

[263] 'Russian bill is copy-and-paste of Germany's hate speech law' *Reporters Without Borders* (Paris, 19 July 2017) <https://rsf.org/en/news/russian-bill-copy-and-paste-germanys-hate-speech-law> accessed 30 October 2018

[264] Alexander Borzenko and Denis Dmitriev, 'Russian lawmakers drafted a new version of their latest lousy idea to regulate social media. But just how bad is it?' *Meduza* (Riga, 9 April 2018) <https://meduza.io/en/cards/russian-

guidelines require from social media networks to delete online hate speech within 24 hours from the moment they get notified by the government[265]. Thus, these countries may now refer to German Network Enforcement Act when being criticized for restricting freedom of expression by western countries.[266]

As being seen as in breach with freedom of expression and international standards, for the time being, it remains only to wait whether and when this Act will be challenged in front of the German Constitutional Court and/or European Court of Justice.

3.3. India: *Shreya Singhal* case-study

Opposite of ECtHR and Germany, which introduced strict liability regime model, India made a step towards different direction. Namely, in the constitutional review of the Information Technology Act (hereinafter: IT Act), the Supreme Court of India read down one of the Sections of the Act in a manner that introduced the broad immunity model for Internet Intermediaries making them liable for third-party content they host only in the case when they explicitly refuse to obey a court or governmental order. At the same time, it struck down the whole Section related to hate speech offences stating that it was too broad and vague and subject to misuse.

lawmakers-drafted-a-new-version-of-their-latest-lousy-idea-to-regulate-social-media-but-just-how-bad-is-it> accessed 30 October 2018

[265] Muthoki Mumo, 'Social media sites to delete hate mongers' accounts in a day' *Daily Nation* (15 July 2017) <https://www.nation.co.ke/business/Social-media-sites-to-delete-hate-mongers-accounts-in-a-day/996-4016026-qo0k2n/index.html> accessed 30 October 2018

[266] Bernhard Rohleder, 'Germany set out to delete hate speech online. Instead, it made things worse' *The Washington Post* (Berlin, 20 February 2018) <https://www.washingtonpost.com/news/theworldpost/wp/2018/02/20/netzdg/?utm_term=.7a50002de4bf> accessed 2 November 2018

3.3.1. Internet Intermediaries' liability and online hate speech regulations prior to the Shreya Singhal case

India is the third biggest country in the world by the number of Internet users with over 300 million users,[267] where hate speech is widespread and usually used by politicians.[268] At the same time, India is a country with the highest number of blocked contents. Just in the first half of 2015, Facebook blocked more than 20,000 pieces of content in India.[269]

Information Technology Act was enacted in 2000 and, in order to provide more safeguard for Internet Intermediaries, it was amended in 2008.[270] Namely, before 2008, IT Act did not set any limits on Internet Intermediaries liability for third-party content, but after the *Avnish Bajaj* case[271] the legislators saw the necessity for introducing rules which would exempt from liability Internet Intermediaries in certain cases. The case concerned Mr. Bajaj, a CEO of Baazee.com, an eBay subsidiary in India. After a Baazee.com user uploaded pornographic video on the website, Mr. Bajaj was prosecuted under IT Act for publishing the prohibited content, even though he did not have an editorial control over it. This prosecution sparked many dissatisfactions among experts[272] and lead to amendment of IT Act and introduction of the safe harbor regime under Section 79A.

[267] Matti Pohjonen and Sahana Udupa, 'Extreme Speech Online: An Anthological Critique of Hate Speech Debates' (2017) 11 International Journal of Communication, 1177
[268] Gargi Chakrabarti and Saahil Dama (n 43)
[269] Christina Medici Scolaro and Jeff Morganteen, 'Facebook blocks more content here than in any other country' (*CNBC*, 13 November 2015) <https://www.cnbc.com/2015/11/13/facebook-blocks-more-content-here-than-any-other-country.html> accessed 17 October 2018
[270] Martin Hvidt Thelle and others, 'Closing the Gap – Indian Online Intermediaries and a Liability System Not Yet Fit for Purpose' (Global Network Initiative, 2014) <https://www.copenhageneconomics.com/dyn/resources/Publication/publicationPDF/1/251/0/Closing%20the%20Gap%20-%20Copenhagen%20Economics_March%202014.pdf> accessed 12 October 2018
[271] *Avnish Bajaj v State of Delhi* (2005) 116 DLT 427
[272] John Ribeiro, 'Experts criticize arrest of Baazee.com CEO' *Macworld* (20 December 2004) <https://www.macworld.com/article/1041530/experts.html> accessed 16 October 2018

Safe harbor regime required from Internet Intermediaries to act only if they had actual knowledge about the unlawfulness of the third-party content they hosted.

As well, in 2009 the Act was amended to introduce the Section 66A with an aim to regulate certain types of cybercrimes.[273] This Section regulated "sending offensive messages" online, in other words it regulated what many would define as hate speech online. Under this Section, a person could face criminal punishment, namely an imprisonment up to three years and a fine if he or she sends online "(a) any information that is grossly offensive or has menacing character; or (b) any information which he knows to be false, but for the purpose of *causing annoyance, inconvenience,* danger, obstruction, insult, injury, criminal intimidation, enmity, hatred or *ill will*, persistently by making use of such computer resource or a communication device; or (c) any electronic mail or electronic mail message for the purpose of causing annoyance or inconvenience or to deceive or to mislead the addressee or recipient about the origin of such messages."[274]

In 2011, the secondary legislation was introduced in India. Namely, under IT Act, Indian government enacted Information Technology Rule, or otherwise called Information Guidelines, whose aim was to give guidelines to Internet Intermediaries how to act in order to comply with IT Act.[275] Primarily aim of the Guidelines enactment was to define what "due diligence" constituted under Section 79(2)(c) in order to prevent Internet Intermediaries self-regulation[276], since this Section required from Internet Intermediaries to "observe due diligence" while they perform their duties under the IT Act. However, it is worth mentioning that these guidelines were heavily criticized due to their vagueness, especially because they did not provide the possibility to have

[273] Suvidutt M. Sundaram and Aditya Tomer (n 13), 146
[274] Information Technology Act (n 20), Sec 66A
[275] Martin Hvidt Thelle and others (n 270)
[276] Gargi Chakrabarti and Saahil Dama (n 43)

the removed information restored, there was no rule about the obligation to inform the third-party provider of information about the takedown, the intermediary was under no obligation to provide a reasoned decision for rejecting or accepting a takedown notice, etc. [277]

After all the amendments to the IT Act, at the time the petitioners brought the case before the Supreme Court about the constitutionality of certain Sections of the Act, liability solution of Internet Intermediaries in India for third-party content, including hate speech, was as follows.

Section 79 of the IT Act, as already mentioned, provided a safe harbor regime for Internet Intermediaries when it came to third-party content they hosted. The Section, *inter alia*, states that Internet Intermediaries will not be held liable for the third-party content they host[278] if they met three conditions. First, if their function is limited to "providing access to a communication system over which information made available by third parties is transmitted or temporarily stored or hosted".[279] Second condition requires from Internet Intermediaries not to "initiate the transmission, select the receiver of the transmission, and select or modify the information contained in the transmission".[280] And finally, the third condition for the exemption from liability prescribe an obligation for Internet Intermediaries to "observe *due diligence* while discharging his duties under the IT Act and also observe such other guidelines as the Central Government may prescribe in this behalf."[281]

However, the IT Act also prescribes the situations in which an Internet Intermediary, although satisfying the requirements set down in Section 79(2), may be held liable for third-party content it

[277] Rishabh Dara, 'Intermediary Liability in India: Chilling Effects on Free Expression on the Internet' (*The Centre for Internet & Society*, 27 April 2012) <https://cis-india.org/internet-governance/chilling-effects-on-free-expression-on-internet> accessed 13 October 2018
[278] Information Technology Act (n 20), Sec 79(1)
[279] ibid, Sec 79(2)(a)
[280] ibid, Sec 79(2)(b)
[281] ibid, Sec 79(2)(c)

hosts. In that manner, Internet Intermediary will be held liable if it "has conspired or abetted or aided or induced, whether by threats or promise or otherwise in the commission of the unlawful act"[282] or if it "upon receiving *actual knowledge*, or on being notified by the appropriate Government or its agency that any information, data or communication link residing in or connected to a computer resource controlled by the intermediary is being used to commit the unlawful act, the intermediary fails to expeditiously remove or disable access to that material on that resource without vitiating the evidence in any manner" [283]. This provision prior to *Shreya Singhal* case, contained a notice-and-takedown procedure, whereby Internet Intermediary was required to take down content upon receiving actual knowledge about its alleged unlawfulness. This, as it will be presented, would change when Supreme Court read down the Section 79(3)(b) and term "actual knowledge" in a manner which would require a court or governmental order specifically requiring from Internet Intermediary to remove certain content.

As regards timeframe in which Internet Intermediary should act in order to avoid being held liable, Information Guidelines in the Rule 3(4) impose obligation to Internet Intermediaries to act *within thirty-six hours* from the moment they obtain knowledge *by itself* or *by an affected person* about unlawful content and work with the user or the owner of such information to disable it.[284] In thirty-six hours upon obtaining the knowledge about unlawful content, Internet Intermediary should just acknowledge or respond to the complainant, while it is obliged to redress the complaint within one month period from the moment the issue was brought up before it.[285] Additionally, the

[282] ibid, Sec 79(3)(a)
[283] ibid, Sec 79(3)(b)
[284] Information Technology (Intermediaries Guidelines) Rules, 2011, Gazette of India, pt III sec 4 (11 April 2011) Rule 3(4)
[285] Ministry of Communications and Information Technology Department of Electronics & Information Technology, 'Clarification on The Information Technology (Intermediary Guidelines) Rules, 2011 under section 79 of the Information Technology Act, 2000' (2013) <http://cyberlawindia.com/wp-content/uploads/2014/06/Clarification-79rules1.pdf> accessed 16 October 2018

intermediary should have a "publicly accessible and published grievance redressal process by which complaints can be lodged."[286] Intermediaries may remove a content which is not compliant with rules and regulations,[287] including the content which is "hateful" or "or otherwise unlawful in any manner whatsoever".[288] Such regulation introduced privately administrated takedown mechanism, where Internet Intermediaries are obliged to remove the allegedly unlawful content upon the complaint of the alleged victim, including the hate speech content as defined in Section 66A.[289]

Under Rule 3(2) of the Information Guidelines Internet Intermediaries are obliged to, observing due diligence, inform their users that they should not post content which is *inter alia* "grossly harmful, hateful, disparaging, otherwise unlawful in any manner whatever".[290]

Additionally, with respect to the Internet Intermediaries' liability regime prior to *Shreya Singhal* case, under Section 69A Internet Intermediary may be requested by the Government to block access to certain content "in the interest of sovereignty and integrity of India, defense of India, security of the State, friendly relations with foreign States or public order or for preventing incitement to the commission of any cognizable offence relating to above".[291] In case it fails to do so, it may subjected to up to seven years imprisonment and a fine.[292] For blocking the content, Government does not need a court order, however, blocking must be performed only when

[286] ibid
[287] ibid
[288] Information Technology (Intermediaries Guidelines) Rules (n 284), Rule 3(2)(b)
[289] Rishabh Dara, 'Intermediary Liability in India: Chilling Effects on Free Expression on the Internet' (2011) <https://cis-india.org/internet-governance/intermediary-liability-in-india.pdf> accessed 13 October 2018
[290] Information Technology (Intermediaries Guidelines) Rules (n 284), Rule 3(2)(b)
[291] Information Technology Act (n 20), Sec 69A
[292] ibid, Sec 69A(3)

absolutely necessary and for the reasons already set down in Article 19(2) of the Indian Constitution. Additionally, reasons for blocking order must be in writing.[293]

From what was mentioned so far, it could be concluded that both the IT Act and its Information Guidelines and EU E-Commerce Directive in similar way regulated Internet Intermediaries liability for third-party content by introducing safe harbor regime and prohibiting general obligation to monitor content they store or transmit. Both Directive and its Guidelines and E-Commerce Directive lack provisions about the possibility to restore the content which was removed due to the err. As well, none of them discusses the possibility for counter-notice – notice to the user that his content has been removed. However, there is an important difference. While E-Commerce Directive makes difference between "mere conduits", "caching" and "hosting" intermediaries and specifies the requirements for each of them, providing safe harbor regime only to the passive intermediaries, IT Act and its Guidelines do not make such difference.

These mentioned Sections, namely Section 66A which regulates hate speech online, Section 79 which regulates Internet Intermediaries' liability for third-party content, and Section 69A which regulates Internet Intermediaries' liability in the cases of online content blocking on the basis of governmental order, were challenged before Supreme Court of India in the following case.

3.3.2. Challenges raised by the petitioners

In the *Shreya Singhal* case[294], petitioners challenged Section 66A of the IT Act claiming it was too broad and vague and as such violated Article 19 of the Indian Constitution which guarantees freedom of expression. Namely, the petitioners emphasized that using terms such as "annoyance", "inconvenience", "danger", "obstruction", "insult", "injury", "criminal intimidation", "enmity",

[293] ibid, Sec 69A(1)
[294] *Shreya Singhal* (n 12)

"hatred" or "ill-will" leads to misuse and suppression of freedom of expression by the authorities[295] and as such are not covered by Article 19(2) which lists exemptions.[296] They emphasized that a person could not be certain which expression would be punishable under the law nor there was a clear guidance given by the authorities how to make a clear line between prohibited and legitimate speech.[297] They also mentioned the possibility of being imprisoned for a term of three years if violating this Section, which additionally caused chilling effect among citizens. Thus, they requested the Section to be declared unconstitutional.

Further claims were related to the Section 79 of the IT Act, where petitioners claimed that it was unconstitutional to force Internet Intermediaries to decide on their own about lawfulness of the third-party content they host, since they represent only a neutral platform where other people interact.[298] At the same time and for the same reason, they challenged the Rule 3(4) of the Information Guidelines, because it was requiring from the Internet Intermediaries to decide on their own whether the third-party content they hosted was illegal and to remove it in the case it was contrary to the Rule 3(2), for which they claimed it was overbroad and vague. Generally, this Section was heavily criticized for its chilling effect and censorship, from different range of reasons, such as: prohibition of hosting content that falls outside of the scope of the Article 19(2) of the Constitution, introduction of private censorship where content could be removed without competent authorities' order, lack of safeguards for the user whose content is removed – the user is not notified about the removal nor the intermediary is obliged to give a reasoned decision, same rules are set for different types of Intermediaries (e.g. registrars for domain names is treated on the

[295] This Section was used many times to suppress political criticism, see: '13 infamous cases in which Section 66A was misused' (*IndiaToday.in*, March 25, 2015) <https://www.indiatoday.in/india/story/section-66a-cases-how-it-curbed-245739-2015-03-24> accessed 18 October 2018
[296] *Shreya Singhal* (n 12), para 5
[297] ibid, para 50
[298] ibid, para 114

equal basis as hosting Internet Intermediary such as social media network) and lack of safeguards which would prevent abuse of take-down notice, since there is no mechanism which would enable once deleted content to be reinstalled.[299]

In the end, the petitioners also claimed that Section 69A should be declared unconstitutional because it did not provide enough safeguards to prevent misuse of blocking orders by government, including lack of the hearing before issuing blocking order.[300] This Section prescribes additional obligation to Intermediaries, as already explained. Namely, it obliges them *inter alia* to block the content they host upon the Government's request for the reasons set down in IT Act.[301] If they fail to do so, they may be subjected up to seven years of imprisonment and a fine.[302] This blocking requirements stand irrespective of Intermediaries' obligation based on Section 79(3) which prescribes exemption from the liability, as already explained.

3.3.3. The Supreme Court's holding and reasoning

In summary, the Court has struck down Section 66A, read down Section 79(3) and upheld Section 69A.

As regards Section 66A, the Court struck the whole section on the ground of vagueness and over-broadness by explaining, as well, the chilling effect the Section had on freedom of expression. The Court backed up its reasoning by referring to the US Supreme court case-law. It stated that Section

[299] Sunil Abraham, 'Shreya Singhal and 66A: A Cup Half Full and Half Empty' (2015) L (15) Economic & Political Weekly 12, 13-14
[300] *Shreya Singhal* (n 12), para 108
[301] The Section 69A(1) of the IT Act reads as follows: "Where the Central Government or any of its officers specially authorized by it in this behalf is satisfied that it is necessary or expedient so to do, in the interest of sovereignty and integrity of India, defense of India, security of the State, friendly relations with foreign States or public order or for preventing incitement to the commission of any cognizable offence relating to above, it may subject to the provisions of sub-section (2), for reasons to be recorded in writing, by order, direct any agency of the Government or intermediary to block for access by the public or cause to be blocked for access by the public any information generated, transmitted, received, stored or hosted in any computer resource."
[302] Information Technology Act (n 20), Sec 69A(3)

66A does not make difference between mere discussion which may be annoying and incitement which can directly lead to public disorder.[303] There is a lack of the link between the content prohibited under Section 66A and the action that may happen as a result of that content.[304] In other words, mere annoyance, causing inconvenience, etc., that are prohibited under Section 66A, as an example, cannot anyhow cause disturbance of public order, as one of the legitimate aims for interference with freedom of expression, neither they can represent an incitement to an offence[305]. Additionally, the Court emphasized that vagueness of this Section resulted from the lack of definitions for expressions used in this Section and the fact that a person, due to the lack of guidance, could not with certainty assess whether a particular speech was punishable or not.[306] It gave an example of two judgements where even courts have reached completely different conclusions on the same facts, depending how they interpreted the term "grossly offensive" or "menacing".[307] Thus, it struck this Section completely in order to prevent further misuse.

As regards Section 79, the Court read down the Section 79(3)(b) of the IT Act and Rule 3(4) of the Guidelines and ruled that Internet Intermediary is obliged to remove content only if there is an order issued by the competent authority only on those grounds set down in Art. 19(2) of the Constitution of India. The court read down the "actual knowledge" term used in this Section and explained that Internet Intermediary is assumed to have actual knowledge about unlawful content upon it receives a court order or being notified by the government or its agency stating that and requiring from it the removal of the impugned content. This means that Internet Intermediaries will be held liable for unlawful third-party content they host only is they disobey the court or

[303] *Shreya Singhal* (n 12), para 20, 35
[304] ibid, para 35
[305] ibid, para 44
[306] ibid, para 76
[307] ibid, para 79-82

government order. In its reasoning, the Court stressed that holding otherwise would make operation of some Intermediaries, as Facebook and Google, impossible, since they would not be able to assess legitimacy of users' requests by their own.[308] By rendering such decision, the Court actually applied Manilla principle which require the possibility to restrict a content only be an order of judicial authority.[309] This solution goes way beyond the regimes installed in Europe. Additionally, the Court emphasized that Internet Intermediary may be obliged to restrict only the unlawful content which falls within Article 19(2) of the Constitution of India. This means that hate speech provisions under Section 66A, which were too broad and rendered unconstitutional, will not apply anymore and previously set standard will be back in use.

Finally, as regards Section 69A, the Court has upheld it thoroughly, emphasizing the fact that Government will request blocking of content only in the cases when blocking is absolutely necessary and for the reasons already set down in Article 19(2) of the Indian Constitution.[310] Additionally, as blocking orders must be recoded in writing, the necessary safeguards are placed, and they can be challenged by writ petitions.[311] Challenging of the orders may take place even before the order is implemented, since the intermediary and the user who posted contested content, if identified, have a right to be heard before the committee.[312] Additionally, the Court emphasized that Government's request for blocking would have to go through Committees scrutiny test, thus, a certain content may be blocked only in the case when Committee decides it is necessary.[313] Thus, the Court did not see this Section as unconstitutional.

[308] ibid, para 117
[309] Manila Principles on Intermediary Liability (n 143), prin 2
[310] *Shreya Singhal* (n 12), para 109
[311] ibid, para 109
[312] ibid, para 110
[313] ibid, para 110

3.3.4. Critiques of the judgement

Some authors claim that the court did not do enough by only reading down the Section 79A. Namely, they emphasize the fact that that Section does not provide safeguards as Section 69A, in the sense that it does not require the existence of a committee which would evaluate the necessity of issuing restriction order.[314] Executive is still entitled to request removal of the third-party content without court order, and since they can issue sanctions for non-compliance[315], Intermediaries, being under pressure, usually comply with their requests.[316] Notwithstanding, Supreme Court's reading down of Section 79 has an important impact on Internet Intermediaries and freedom of expression, since they will be obliged to act only upon receiving an official court or governmental order, and not a private party. As such, no private administrative censorship will take place.

As regards blocking orders issued by Government under Section 69A, it is not contested that Government, in order to prevent riots, may issue orders for blocking content that spread hatred and call for violence. However, Government may often misuse their power, since the link between online content and riots in many cases does not exist.[317] Thus, some authors emphasize that an executive order for blocking content must fulfil additional requirement, namely, it must answer to the question whether there is a direct link between online content and violence.[318] Additional problematic part of these Section which is not addressed is the possibility to impose a criminal

[314] Jyoti Panday, 'The Supreme Court Judgment in Shreya Singhal and What It Does for Intermediary Liability in India?' (*The Centre for Internet & Society*, 11 April 2015) <https://cis-india.org/internet-governance/blog/sc-judgment-in-shreya-singhal-what-it-means-for-intermediary-liability> accessed 15 October 2018

[315] In 2012 during the Assam violence in India, Government threatened to ban Twitter operating in India if they failed to remove the user pages that contained hate speech.

[316] Gargi Chakrabarti and Saahil Dama (n 43)

[317] A study done by LSE and The Guardian regarding the Tottenham riots showed that use of social media did not enhanced violence. See: The Guardian and LSE, 'Reading the Riots Investigating England's Summer of Disorder' (2011) <http://eprints.lse.ac.uk/46297/1/Reading%20the%20riots(published).pdf> accessed 18 October 2018

[318] Gargi Chakrabarti and Saahil Dama (n 43)

punishment for failing to comply with government order. It it quite disputable to what extent such solution is indeed proportionate.

At the end, some authors claim that the decision does not talk about necessity for establishing the proper safeguards such as a prior hearing before issuing the restriction (under Section 79A) or blocking order (under Section 69A), nor there is any other opportunity to appeal the order except a writ petition.[319] Additionally, no limitation are placed upon on length, duration and geographical scope of the content restriction, and nothing was mentioned about the lack of possibility for the third-party user whose content has been removed to restore it.[320]

Aside from the critiques, the judgement is seen as a victory for freedom of expression which stroke down chilling effect, since legality of the content is now to be decided by the competent authorities and not private parties.[321]

B. Conclusion

Internet Intermediaries may include many different types, but for the purpose of this thesis the emphasis was put on the ones that host third-party content without anyhow modifying it. Generally speaking, these Internet Intermediaries may or may not be found liable for the unlawful content they host, depending on the type of liability model adopted in a jurisdiction where they operate. Nowadays, there is an increase in the introduction of strict liability regimes in the cases of online hate speech, which consists of imposing a monitoring obligation to Internet Intermediaries. This solution was adopted by the ECtHR after *Delfi AS* judgement and Germany after the enactment of

[319] Jyoti Panday (n 314)
[320] ibid
[321] Kartik Chawla, 'Shreya Singhal, and How Intermediaries are Simply Intermediaries Once Again – Striking Down the Chilling Effect' (*Tech Law Forum*, 30 March 2015) <https://techlawforum.wordpress.com/2015/03/30/shreya-singhal-and-how-intermediaries-are-simply-intermediaries-once-again-striking-down-the-chilling-effect/> accessed 13 October 2018

the Network Enforcement Act and it has been heavily criticized since then as being detrimental to freedom of expression. On the other hand, some countries, like India, shifted away from such solutions, offering a broad immunity model to Internet Intermediaries for online hate speech they host, by prescribing they would be held liable only and exceptionally in cases they fail to obey a court or governmental order.

CHAPTER III STRICT LIABILITY REGIME – IMPLICATIONS AND REASONS BEHIND ITS IMPOSITION

A. Introduction

As shown, strict liability regime has been introduced in Europe, where EU leads a campaign influencing Internet Intermediaries to take firm steps and start monitoring and deleting harmful content. This model has been heavily criticized due to the implications it has on human rights. Thus, this Chapter will start with addressing all of the negative implications following strict liability regime, taking into account the jurisdictions explained in the previous Chapter. It will especially put emphasize on the lack of judicial protection of freedom of expression and high level of censorship introduced as a result of the obligation to monitor content they host. At the end, the Chapter will address some of the possible reasons and tendencies why the Europe has started imposing this regime.

1. Implications of strict Internet Intermediaries' liability in cases of online hate speech

Strict liability regime raises many negative implications for human rights, especially freedom of expression. This regime, as already elaborated, results in pre-monitoring and filtering requests, which are "capable of undermining freedom of the right to impart information on the Internet"[322]. It represents a threat to open Internet, public debate and exchange of ideas.[323] These negative implications resulting from strict liability regime are even more highlighted in the situation when monitoring and filtering is required for dealing with the content which may be regarded as hate speech, due to its contested nature and lack of unanimous definition.

[322] *Magyar Tartalomszolgáltatók Egyesülete and Index.Hu Zrt* (n 208)
[323] UNHRC (n 16), paras 44-48

Intermediaries should never be obliged to generally monitor all the content they host.[324] Imposing strict liability regime to Internet Intermediaries in cases of online hate speech, as in the case of ECtHR and Germany, bears many detrimental consequences, which will be elaborated as follows.

1.1. Lack of judicial protection of freedom of expression

The decision whether some content indeed represents a hate speech and as such is suitable for removal should be a matter of judicial authority, since that is a legal question. Intermediaries, as private parties, should not be placed in a position to decide about the legality of third-party content.[325] They should not be obliged to remove certain content unless a judicial authority determined the unlawfulness of the content[326] or other impartial, independent and authoritative oversight body.[327] As noted by UN Special Rapporteur on Freedom of Opinion and Expression, the OSCE Representative on Freedom of the Media and the OAS Special Rapporteur on Freedom of Expression in their 2005 Joint Declaration, "no one should be liable for content on the Internet of which they are not the author, unless they have either adopted that content as their own or refused to obey a court order to remove that content."[328]

[324] Manila Principles on Intermediary Liability (n 143), prin 1(b), 1(d)
[325] ibid, prin 3(a); UNGA (n 28), para 87; Joint Declaration on the Internet and on Anti-Terrorism Measures by the UN Special Rapporteur on Freedom of Opinion and Expression, the OSCE Representative on Freedom of the Media and the OAS Special Rapporteur on Freedom of Expression (2005); Article 19 (n 6)
[325] Joint Declaration on the Internet and On Anti-Terrorism Measures by the UN Special Rapporteur on Freedom of Opinion and Expression, the OSCE Representative on Freedom of the Media and the OAS Special Rapporteur on Freedom of Expression (n 325)
[326] Manila Principles on Intermediary Liability (n 143), prin 2(a)
[327] Joint Declaration on Freedom of Expression and "Fake News", Disinformation and Propaganda by the United Nations Special Rapporteur on Freedom of Opinion and Expression, the Organization for Security and Co-operation in Europe Representative on Freedom of the Media, the Organization of American States Special Rapporteur on Freedom of Expression and the African Commission on Human and Peoples' Rights Special Rapporteur on Freedom of Expression and Access to Information (2017)
[328] Joint Declaration on the Internet and On Anti-Terrorism Measures by the UN Special Rapporteur on Freedom of Opinion and Expression, the OSCE Representative on Freedom of the Media and the OAS Special Rapporteur on Freedom of Expression (n 325)

Hate speech, as already shown in first Chapter, is not universally accepted and defined. Due to the thin and not clearly established line between the hate speech and freedom of expression, and due to the lack of the answer to the question what hate speech actually entails, it is very hard to make difference between clearly unlawful content and lawful one even for judges. For that reason, only a mechanism which would enable an oversight by judicial body would be in accordance with international human rights law.[329] In that manner, Article 19 emphasize that in the case of incitement to hatred and violence, law enforcement agencies or judicial authorities must be contacted.[330] This may not apply only in the cases when the content is illegal notwithstanding the context, such as in the case of child pornography.[331]

The Governments' claims that Internet Intermediaries are better situated to address the unlawful content due to the better technical means they have, does not stand, since technical means are not the important factor, but judicial assessment of the content.[332] When the power to decide about the lawfulness of the content is fully vested in Internet Intermediaries, without proper judicial examination, it may happen that Internet Intermediaries, out of caution, make a mistake and forbid legitimate content, which is more probable in the cases of hate speech.[333] Apart from this, it may happen that they use their position and delete content they do not like.

Solutions adopted in ECtHR and Germany vest the power to decide upon legality of the content that allegedly represents hate speech to private entities that lack legal knowledge. As already mentioned, decision about whether a certain content amounts to hate speech is a tough task even

[329] Opinion of David Kaye Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression regarding the German Draft Law Netzdurchführungsgeset (n 234)
[330] Article 19 (n 6), p 18
[331] Council of Europe Committee of Ministers (n 152), p 1.3.2.
[332] Article 19 (n 6), p 14
[333] Wolfgang Benedek and Matthias C. Kettemann, *Freedom of expression and the Internet* (Council of Europe Publishing 2013), 104; Office of the Special Rapporteur for Freedom of Expression of the Inter-American Commission on Human Rights (n 15), para 105

for courts, including the ECtHR itself.[334] Lawfulness of the content is something that should be objectively and not subjectively tested.[335]

In the cases of both ECtHR's case *Delfi AS* and German Network Enforcement Act, except obliging the Intermediaries to decide upon the lawfulness of the content they host, no clear guideline is given to help them decide what "unlawful content" actually entails, which additionally runs counter freedom of expression.[336] Although the Network Enforcement Act lists down offences which fall under unlawful content, the problem arises with the fact that the list is too broad. As such, it includes offences which need different level of protection, and thus, careful assessment of the facts which would require taking into account the context of the offence. For example, defamation of religions, religious and ideological associations will not require the same level of protection as preparation of a serious violent offence endangering the state. For these reasons, Internet Intermediaries, as private actors, would not be capable of assessing all the circumstances needed and rendering the best decision.

In that sense, India, as one of the analyzed jurisdictions, made a step forward towards protection of freedom of expression, by leaving the question of lawfulness of the content to the hands of court. However, the solution given in India raises a bit different implication with respect to the power of the executive to order the removal of third-party content apart from the judicial one, which were mentioned in the previous Chapter.

[334] Often there are divided votes on this issue; see, *inter alia, Vejdeland and others v Sweden* App no 1813/07 (ECtHR, 9 February 2012), *Féret* (n 38), *Perinçek* (n 21)
[335] Rishabh Dara (n 277)
[336] Dirk Voorhoof and Eva Lievens, 'ECtHR confirms and tempers Delfi judgment: operators of Internet portals not liable for dissemination of offending - but not "clearly unlawful" - user comments' (*ECHR BLOG*, 15 February 2016) <http://echrblog.blogspot.com/2016/02/offensive-online-comments-new-ecthr.html> accessed 11 October 2018

1.2. Censorship

One of the most problematic consequence of the introduction of strict liability regime is censorship. Monitoring user-generated content which follows from strict liability regime is equivalent to endorsing a form of private censorship.[337] Introducing a strict liability regime which would require Intermediaries to monitor content does not mean that censorship would be directly imposed by the Governments. Rendering a decision such as *Delfi AS* and adopting a law such as Germany did, impose so-called "collateral censorship"[338]. Governments do not directly censor expression online, but they put pressure on Internet Intermediaries, as the ones who enable the exchange of ideas and increase public discourse online. Pressure comes from the fear of liability which would Intermediaries encounter in case they fail to comply with the requirements set by laws adopted by the Government, as in Germany[339], or by regional courts binding rulings, as in ECtHR practice. In that sense, both ECtHR and Germany expect Intermediaries to monitor the content they host and expeditiously delete it if it represents hate speech. German law additionally prescribes huge fines, which taken together with short time period in which Internet Intermediaries are obliged to remove content, undoubtedly leads to introduction of self-censorship, or otherwise named "pre-cautionary" censorship,[340] since as explained, Internet Intermediaries might delete the content out of caution in order to avoid liability.

Strict liability regime may lead to suppression of users' speech, as they may refrain from posting comments if their comments are being heavily and unjustifiably removed by Internet Intermediaries. Additionally, it may lead to the suppression of users' speech in the sense that

[337] UNHRC (n 16), para 40; Jack M Balkin, 'Old-school/New-school Speech Regulation' (2014) 127 Harvard Law Review 2296, 2309; Cynthia Wong and James X Dempsey (n 17)
[338] *Delfi AS* (n 10) dissenting opinion of judges Sajo and Tsotsoria, para 2
[339] BIU – Bundesverband Interaktive Unterhaltungssoftware (n 227)
[340] Opinion of David Kaye Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression regarding the German Draft Law Netzdurchführungsgeset (n 234)

Internet Intermediaries, due to the huge fines they may face and by not being able to cope with the obligation to monitor all the content they host and delete it within extremely short period of time, start deleting legal content from their platforms without conducting a proper and thorough investigation whether such content is indeed a harmful one.[341] As well, they may shut down their business completely or they might stop offering a forum for hosting comments and put an end to public debate[342]. This is a natural consequence since Internet Intermediaries' interest in not being held liable will prevail over their interest to protect freedom of expression online.

Opposite to the strict liability regime and solutions adopted in ECtHR and Germany, novelties that were brought into Indian legal system after rendering a decision in the case of *Shreya Singhal*, especially the read down of Section 79(3)(b) of the IT Act, whereby Internet Intermediary is held liable only if it fails to obey a court or governmental order requiring a removal of unlawful content, are seen as limits on private censorship.[343] Prior to reaching such decision, Internet Intermediary was required to delete content upon receiving actual knowledge within one month period.[344] Even without the imposition of the obligation to perform general monitoring, this regime whereby an Intermediary is obliged to remove unlawful content without actually obtaining a court order, was regarded as "state-mandated private censorship regime".[345]

As a proof for existence of high level of censorship when it comes to the removal of the content done by Internet Intermediaries, a study was done in India prior to the ruling in Shreya Singhal case. The study has shown that Internet Intermediaries tend to take down content upon notification

[341] *Delfi AS* (n 10), joint dissenting opinion of judges Sajo and Tsotsoria, para 8
[342] Ibid, join dissenting opinion of judges Sajo and Tsotsoria, para 1
[343] Jyoti Panday (n 314)
[344] Ministry of Communications and Information Technology Department of Electronics & Information Technology (n 285)
[345] Gargi Chakrabarti and Saahil Dama (n 43)

automatically, despite the apparent flaws without making a proper assessment of its illegality.[346] For example, one of the seven Internet Intermediaries which were part of the study, was an online news portal which published an article about the Telengana movement, a movement consisting of political activists who wanted to create a new state in south India. Below the article there was a comment box and one of the user's wrote a comment in which he condemned the violence which was advocated by Telengana leaders and addressed them different questions about the movement. After the conducer of the study reported the comment *inter alia* as hateful and requested its removal, the Intermediary removed not only that content, but all the others that were posted below the article, 15 in total, whereby most of them were completely lawful.[347] In this manner, the study showed that Internet Intermediaries often comply mechanically with removal requests, even when content is clearly lawful. It is not necessary to mention that this level of censorship would be undoubtedly even more enhanced in the case when they are required to act expeditiously if they do not want to face huge fines or criminal punishment for failing to remove allegedly harmful content.

1.3. Unreasonable expectations from Internet Intermediaries and lack of resources to comply with imposed obligations

Internet intermediaries that host significant number of third-party contents, may not have a possibility to comply with requirements which strict liability regime imposes. Monitoring of a content in order to prevent appearing of the allegedly hateful third-party content, requires additional resources and enormous personnel in order to pre-screen entire user-generated content.[348] Only Facebook hired nearly double number of employees around the world to enhance

[346] Rishabh Dara (n 289)
[347] Among the deleted content there was e.g. simple question whether students of certain University were involved in the movement and advocated for violence.
[348] Cynthia Wong and James X Dempsey (n 17), 14

removal of allegedly unlawful content.[349] But it is disputable how can even a hiring of employees represent a good solution, since too many people would be needed to moderate content and go through all complaints. Even if hired, the same question will remain: How can those people decide what content to delete?

Imposing such obligation to Internet Intermediaries can have an inhibitory effect on them, which could eventually influence their future operation and eventual shut down.[350] Making Internet Intermediaries liable for third-party comments under strict liability regime represents an unacceptable burden,[351] which alters their functionality[352]. Popular social media might not survive if they are not offered with potential liability shield.

Internet Intermediaries, both by German law and ECtHR jurisprudence, are put in the situation where they may face fines if they do not delete content which represent hate speech, although they lack clear guidelines about what hate speech is. This is especially problem in Germany, where fines are enormous, reaching up to 50 million EUR. Thus, Internet Intermediaries may face economic consequences through payment of fines if they do not delete enough posts, or, on the other hand, they may face the same consequences if they, to avoid liability, start posting content which is not related to questions of public interest which would trigger public debate or disable the opportunity for hosting third-party content[353], and by doing so, lose their customers. At the end, this situation may also influence other Internet Intermediaries not to even start their business, thus limiting as well competition on the market.[354]

[349] Melissa Eddy and Mark Scott (n 11)
[350] Cynthia Wong and James X Dempsey (n 17), 14
[351] Marketa Trimble and Salil K. Mehra, 'Secondary Liability, ISP Immunity, and Incumbent Entrenchment' (2014) 62 Am J Comp L 685, 700, *Delfi AS* (n 10), para 96
[352] Case 8611/12 (n 152)
[353] Dirk Voorhoof and Eva Lievens (n 336)
[354] EDRi (n 230)

1.4. Lack of protection for third-party rights

First of all, imposition of strict liability regime, where Internet Intermediaries are expected to perform the role of the judge and decide expeditiously upon the unlawfulness of the third-parties' content breaches their due process rights. In other words, the situation undermines due process rights of the third-party whose comments are deleted, since no private party should decide about lawfulness of their posts.

Additionally, in the cases when Internet Intermediaries remove the third-party content, an effective redress mechanism should be put in place for third-parties.[355] However, third parties usually cannot appeal the decision for taking down their content, which is in breach of their due process rights.[356] In that manner, the ECtHR does not provide any procedural guarantees[357], nor does the German Network Enforcement Act. Namely, in Germany, under the Network Enforcement Act, although there is an Internet Intermediary's obligation to inform third-party about its decision to take down the content due to its alleged unlawfulness[358], there is no right to appeal the decision in the case lawful content is illegitimately blocked or removed from the platform.[359] In certain situations, provisions from national contract laws can serve as a source for third-parties to claim damages and violation of their rights. Third-parties could call upon breach of the contract in a situation when terms of services between the third-party and the Internet Intermediary are violated.[360] However, this may not be regarded as the best solution.

[355] Council of Europe Committee of Ministers (n 152), p 1.3.3.
[356] Manila Principles on Intermediary Liability (n 143), prin 5(b)
[357] Dirk Voorhoof and Eva Lievens (n 336)
[358] NetzDG Art 1 Sec 3(2)(5)
[359] Article 19 (n 235) p 17
[360] Divij Joshi (n 144), p 24

As well, third-party should have an opportunity to be heard about the legality of their content and request for its removal, if not before the restriction of the content, then as soon as possible after the restriction.[361] Additionally, there should exist the opportunity for a third-party content that proves to be lawful to be reinstalled.[362] Lack of the provision which would prescribe the re-installment of the content that was initially taken down by the Intermediaries can lead to irreversible harm for the third-parties.[363]

At the end, Internet Intermediaries under the strict liability regime do not have a duty to give reasons why did they delete certain content. User, whose content is deleted, could probably initiate civil proceedings for breach of his freedom of expression right. However, this shows disproportionate burden that was put on the user, whereas it comes to potential victims, Internet Intermediaries can easily delete content.

Except for due process rights, strict liability regime breaches users' right to privacy and right to impart and receive information. Namely, filtering and monitoring are deeply invasive and represent a very heavy burden for Internet Intermediaries.[364] Filtering involves the processing of IP addresses, which represent personal data. Even in the case of *Sabam v Scarlet,* the CJEU ruled that Internet Intermediaries cannot be imposed with such an obligation as to filter the content, since that would be in breach with users' right to privacy and right to impart and receive information, but also with ISPs rights, such as property right.[365] In that case, which dealt with copyright issues, the Court stressed that intellectual property is not absolutely protected and should be balanced with

[361] Manila Principles on Intermediary Liability (n 143), prin 5(a)
[362] ibid, prin 5(d)
[363] Gargi Chakrabarti and Saahil Dama (n 43)
[364] Merit Ulvik and Darian Pavli, 'Case Watch: A Strasbourg Setback for Freedom of Expression in Europe' (*Open Society Foundations,* 22 October 2013) <https://www.opensocietyfoundations.org/voices/case-watch-strasbourg-setback-freedom-expression-europe> accessed 11 October 2018
[365] Case C-70/10 (n 154), para 44-50

mentioned users' and ISPs rights.[366] In the same manner, this applies to the question of online hate speech and protection of the rights of potential victims which must be balanced with right to privacy and freedom of expression of the users.

1.5. Decrease of public debate through introduction of "real name policy"

One of the consequences that strict liability regime may bring is introduction of "real name policy". In that manner, in the *Delfi AS v Estonia*, the ECtHR concluded that the interference with the Applicant's right to freedom of expression was necessary, since it did not install any measure which would enable potential victims to identify the perpetrators and bring a claim. One of the possible measures to comply with the ECtHR request would be the introduction of real name policies.

Introduction of real-name policy is especially important question for different online forums and online media, where a lot of users post comments anonymously. Introducing a real-name policy may have two consequences. First one is that it would constitute an interference with the right to privacy and freedom of expression of the users, impairing their essence and leading to users commenting less, in other words it would censor the speech. On a long-term plan, it would suppress the public debate. The second consequence is related to the implications that would be faced by Internet Intermediaries and which result from the previously mentioned consequence. As already mentioned, such policy would make users comment less or search for alternative Internet Intermediaries which allow anonymous comments or usage of pseudonyms. This could be especially discriminatory for national Intermediaries, who may even shut down their business in a

[366] ibid, para 43

response, since many users could search for international intermediaries that would offer possibility to stay anonymous, which was the case in Korea.[367]

2. Reasons behind the imposition of strict liability model

What could be the possible reasons for such reckless moves which endanger the foundations of a democratic society as the introduction of strict liability regimes for Internet Intermediaries in the cases of online hate speech? It seems like the world has become particularly worried about its security after technological innovations took place. Rapid developments of Internet technologies cannot be followed by ordinary citizenry, and Internet is still perceived as insufficiently explored place subjected to further improvements. The most important task is placed before legislators who are trying to keep up with all the developments and regulate Internet as much as possible.

Given to the "specific" features of the Internet, such as lack of boundaries, easier visibility and accessibility of online content, itinerancy of online content, possibility of perpetrators to stay anonymous and easily advocate for their causes, etc., people increased their fear from becoming an easy victim of online attack. All around the literature, the emphasis is put on the fact that online crime is more severe than offline one.[368] This type of stance plays a crucial role in assessing the reasons for overreacting moves from ECtHR and Germany when addressing the questions of suppressing online hate speech.

The two main reasons for introducing the strict liability regime to Internet Intermediaries in the cases of online hate speech are primarily irrational fear from being victimized and Governments' need to provide security and welfare to its citizenry.

[367] *Delfi AS* (n 10), para 98-100
[368] UNGA (n 28)

3.1. Irrational fear

Irrational fear comes from even minimal chance of becoming a victim - and due to the nature of Internet, as already mentioned, a person can easily encounter harmful content which increase his or her chances to become a victim, especially in the cases of hate speech, if hate speech is understood too broadly. In other words, if understanding of the hate speech is too broad, people could regard themselves as victims, even in cases the content is not directed towards them. As already mentioned, in the most cases of hate speech the victim cannot be identifiable, since the speech can be directed towards the group in general. This fear has started running counter the whole idea of human rights, by diminishing them.

3.2. Welfare state and right to security

Additional reason for reckless imposition of strict liability regimes to Internet Intermediaries can be found in the idea of the state as the protector of the nation. It seems like the world is shifting back to the time where national states were seen as the only ones who could have provided welfare and security for their nation. The world began to panic, since it remains incapable of dealing with all the challenges Internet has brought. Everyone sees himself as potential victim of online hateful content, and the state is seen as the one that needs to provide absolute security to its citizens and stay focused on preventing and removing any external threat, no matter where it originates from.[369] Thus, the State can easily be influenced by public opinion and pushed to offer higher security protection to their citizenry, which can lead to reaching unbalanced decisions. The states feel they have a duty to protect victims of hate speech in every possible manner.[370]

[369] Simon Hallsworth and John Lea, 'Reconstructing Leviathan: Emerging contours of the security state' (2011) 15(2) Theoretical Criminology 141, 142

[370] Answers to the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression in regard to the Act to Improve Enforcement of the Law in Social Networks (Network Enforcement Act), provided by the Federal Government of Germany (n 221)

The problem comes from the fact that online users who post content are usually anonymous and State cannot track them down, since even when IP addresses can be revealed, exact person who did use the address cannot be found. This may happen due to the fact that Internet was accessed on some public computer, or that maybe the perpetrator used technological means to hide its track. In such situation, State may feel obliged to act as the protector of its citizenry and thus, reach a decision to impose strict liability regime to Internet Intermediaries, making them liable for third-party content even in cases they act expeditiously in removing the allegedly hateful content upon notification, as seen in the ECtHR case.

Consequently, there has been an increase in the opinion that people enjoy right to security which can be balanced with other human rights. As a result, states became driven by security reasons and started overriding respect from core human rights. However, this can be extremely dangerous, since once the state sets security as a constitutional value and starts using it as factual basis of judicial decisions, the risk of preserving judicial state gets higher, and traditional balancing of principles becomes impossible.[371] Thus, relying on the so-called "right to security" indeed represents a bad idea which could only result into demise of human rights, as seen from this example, but a bad idea which is unfortunately emerging.

Disregard the reasons, it seems like Europe started forgetting that freedom of expression is the essential foundation of a democratic society.[372] It is quite shocking how governments, eager to provide the sense of security and peace to their citizens, fail to take a look on the both sides of the coin. Maybe one of the additional problems, lays down in idea of seeing human rights as primarily those rights to which people are entitled since they serve their basic needs, while other rights as

[371] Andras Sajo, 'Symposium: Terrorism, Globalization and the Rule of Law: From Militant Democracy to the Preventive State?' (2006) 27 Cardozo L Rev 2255, 2292
[372] *Handyside* (n 33), para 49

freedom of expression, in the eyes of many traditionalist remain to be of lesser value.[373] Since widespread fear from being victimized easier by use of Internet as a tool, and freedom of expression is seen as less valuable right in comparison to the right to security and possibility to become a victim, a state, as the one whose task is to provide security and welfare, easier allows measures which disproportionately demise freedom of expression, such as introducing the strict liability regime for Internet Intermediaries.

These reckless moves taken by ECtHR and Germany – consisting of imposing strict liability regime to Internet Intermediaries for online hate speech content they host – can cause enormous harm to the protection of freedom of expression and preservation of a democratic society by diminishing their values. There is a need for other states to refrain from following steps of ECtHR and Germany and pay attention not to exaggerate when it comes to assessing the real threats of Internet, rationality of the fear from being victimized, and need for providing security. Without proper assessment of these elements, a serious risk of diminishing human rights prevails. Internet Intermediaries are extremely beneficial not only because they offer channels for dissemination of media content[374], but also because they enhance enjoyment of Internet users' rights to freedom of expression and access to information[375] and this should be borne in mind when assessing their liability.

B. Conclusion

Strict liability regime, as applied in ECtHR and Germany, but in general as well, runs counter international standards. Prescribing an obligation to Internet Intermediaries to monitor all the

[373] Andras Sajo, 'The Fate of Human Rights in Indifferent Societies' (Demise of Constitutionalism Conference, Budapest, May 2018)
[374] Council of Europe Committee of Ministers, 'Recommendation to member states on a new notion of media' (2011) CM/Rec(2011)7, para 63
[375] UNHRC (n 16), para 74

content they host undermines freedom of expression and introduces high level of censorship. At the same time, it represents an excessive burden for Intermediaries, which requires additional resources and enormous personnel in order to pre-screen entire third-party content. It endangers third-parties due process rights and privacy rights. States, driven by wrong reasons, often fail to properly assess the real threats of Internet, rationality of the people's fear from being victimized and need for providing security, thus, introducing strict liability regime, which is undoubtedly detrimental to human rights, as seen in the cases of ECtHR and Germany.

CONCLUSION

Internet technology area has brought up many benefits for the mankind. Internet has been used for accessing different information, sharing views and exchanging ideas and opinions, and as such has contributed to the wide-open debates on different public issues. At the same time, it has been used as a forum for incitement to violence and hatred.

In order to properly address the question of online hate speech, international community has tried to find adequate solution, and as a result, a trend to include Internet Intermediaries into online hate speech suppression has developed. Internet Intermediaries, as the ones who host the third-party content, are starting to be seen by many as crucial players in the hate speech suppression since they allegedly have enough financial and technical means to block or remove hateful content.

As the thesis has shown, the models for Internet Intermediaries' liability in hate speech suppression and generally can differ, ranging from absolute immunity to strict liability. However, on a global level, imposition of strict liability regimes for third-party content posted on Internet Intermediaries' platform has become a trend. Strict liability solution, as shown, requires monitoring of all the content Internet Intermediaries host in order to be able to act expeditiously enough in deleting the one which represent hate speech and avoiding liability and huge fine. Usually, these solutions were defended by emphasizing the seriousness and consequences which online hate speech may entail.

The thesis has shown that, strikingly, Europe started moving towards imposition of strict liability regime, as well, after ECtHR reached a decision that Internet Intermediaries can be held liable if they fail to expeditiously delete content that amounts to hate speech and is hosted on their platform, and after Germany adopted a law which imposed an extremely strict obligation to Internet Intermediaries to take down harmful online content within 24 hours if they do not want to pay

huge fines. In that manner, the thesis has analyzed the strict liability regime introduced with ECtHR's ruling in the case of *Delfi AS* and German Network Enforcement Act. Additionally, it analyzed the solution adopted in India, which completely differed from the ones from Europe, since it moved away from safe harbor regime to broad immunity regime by prescribing that Internet Intermediary may be held liable for third-party hate speech content only in case it fails to obey a court or governmental order. By analyzing these three jurisdictions, although focusing on strict liability regime, the thesis has shown different solution to the same problem – online hate speech suppression – by comparing their impact on freedom of expression.

The thesis has reached to the conclusion that by introducing strict liability regime for online hate speech, countries have failed to protect freedom of expression. The fact that Internet Intermediaries only host the content without anyhow modifying it, must be taken into account when introducing the laws which prescribe their liability. Monitoring resulting from strict liability regime contravenes international standards on liability of intermediaries[376]. Additionally, imposition of strict liability regime in cases of online hate speech reveals many negative implications to human rights, especially freedom of expression as the core right in a democratic society. Strict liability regime undermines freedom of expression and introduces high level of censorship, especially given the fact that Internet Intermediaries would often err and delete content out of caution, which is more probable when the content is allegedly amounting to hate speech, since the illegality of the content is hard to establish even by Courts who do possess required legal knowledge. At the same time, strict liability regime represents an excessive burden for Intermediaries, which requires additional resources and enormous personnel in order to pre-screen entire third-party content.

[376] Council Directive 2000/31/EC (n 127), Art 15; Council of Europe Declaration on freedom of communication on the Internet (n 152), prin 6; Case 8611/12 (n 152); Manila Principles on Intermediary Liability (n 143), prin 1(d); Council of Europe Committee of Ministers (n 152), p 1.3.5.

Additionally, it endangers third-parties due process rights and privacy rights and it decreases public debate.

Finally, the thesis has led to a conclusion that the solution for online hate speech suppression must not be sought by introducing strict liability regime. There is a need for other states to refrain from following steps of ECtHR and Germany and pay attention not to exaggerate when it comes to assessing the real threats of Internet, rationality of the fear from being victimized, and need for providing security. Without proper assessment of these elements, a serious risk of diminishing human rights prevails.

To sum up, the thesis has shown how important Internet and freedom of expression are, emphasizing that the virtues of the Internet should not be used against it. As EU digital Commissioner said: "We shouldn't kill innovation in Europe by over-regulating platforms".[377]

[377] 'EU digital Commissioner attacks German's hate speech bill' *Financial Times* (London) <https://www.ft.com/content/1407bcd8-0d68-11e7-b030-768954394623> accessed 29 June 2018

BIBLIOGRAPHY

Books and Articles

Abraham S, 'Shreya Singhal and 66A: A Cup Half Full and Half Empty' (2015) L (15) Economic & Political Weekly 12

Angelopoulos C and Smet S, 'Notice-and-fair-balance: how to reach a compromise between fundamental rights in European intermediary liability' (2016) 8(2) Journal of Media Law 266

Bakalis C, 'Rethinking cyberhate laws' (2018) 27 (1) Information & Communications Technology Law 86

Balkin M J, 'Old-school/New-school Speech Regulation' (2014) 127 Harvard Law Review 2296

Banks J, 'European Regulation of Cross-Border Hate Speech in Cyberspace: The Limits of Legislation' (2011) 19 European Journal of Crime, Criminal Law and Criminal Justice 1

Benedek W, Kettemann C M, *Freedom of expression and the Internet* (Council of Europe Publishing 2013)

Buffa F, *Freedom of expression in the Internet society* (Key, 2016)

Buyse A, 'Words of Violence: "Fear Speech," or How Violent Conflict Escalation Relates to the Freedom of Expression' (2014) 36 (4) Human Rights Quarterly 779

Cohen-Almagor R, *Speech, Media and Ethics, The Limits of Free Expression: Critical Studies on Freedom of Expression, Freedom of the Press and the Public's Right to Know* (Palgrave 2001)

Cotter T F, 'Some Observations on the Law and Economics of Intermediaries' (2006) 1 Michigan State Law Review 67

Cucereanu D, *Aspects of Regulating Freedom of Expression on Internet* (1st edn, Intersentia, 2008)

Hallsworth S and Lea J, 'Reconstructing Leviathan: Emerging contours of the security state' (2011) 15(2) Theoretical Criminology 141

Harvey D, *Collisions in the Digital Paradigm: Law and Rule-making in the Internet Age* (Hart Publishing 2017)

Hawdon J, Oksanen A and Räsänen P, 'Exposure to Online Hate in Four Nations: A Cross-National Consideration' (2017) 38 (3) Deviant Behavior 254

Mill S J, *On Liberty* (first published 1859, Batoche Books 2001)

Pohjonen M and Udupa S, 'Extreme Speech Online: An Anthological Critique of Hate Speech Debates' (2017) 11 International Journal of Communication 1177

Pollicino O, Bassini M, 'Free speech, defamation and the limits to freedom of expression in the EU: a comparative analysis' in A. Savin and J. Trzaskowski, *Research Handbook on EU Internet Law* (eds) (2014)

Reed C, 'The Challenge of Hate Speech Online' (2009) 18 (2) Info & Comm Tech L 79

Sajo A, 'Symposium: Terrorism, Globalization and the Rule of Law: From Militant Democracy to the Preventive State?' (2006) 27 Cardozo L Rev 2255

Spindler G, 'Internet Intermediary Liability Reloaded – The New German Act on Responsibility of Social Networks and its (In-) Compatibility with European Law' (2017) 8 JIPITEC 166

Sundaram M S and Tomer A, 'Cyberhate in India – Regulation and Intermediary Liability' (2017) 4 (3) International Journal of Law and Legal Jurisprudence Studies 143

Tao Q, 'The Knowledge Standard for the Internet Intermediary Liability in China' (2012) 20(1) International Journal of Law and Information Technology 1

Trimble M, Mehra K S, 'Secondary Liability, ISP Immunity, and Incumbent Entrenchment' (2014) 62 Am J Comp L 685

Weber R H, 'Challenges for Communications in a Changing Legal Landscape' in D. Weisenhaus and S. NM Young (eds), *Media Law and Policy in the Internet Age* (Hart Publishing, 2017)

Weckert J, 'What is so bad about Internet content regulation?' (2000) 2 Ethics and Information Technology 105

Wong C and Dempsey J X, *Mapping Digital Media: The Media and Liability for Content on the Internet* (Open Society Foundations, 2011)

Case law material

"The Last Temptation of Christ" (Olmedo-Bustos et al) v Chile, Inter-American Court of Human Rights Series C No 73 (5 February 2001)

Abrams v US 250 US 616 (1919)

Ahmet Yildirim v Turkey App no 3111/10 (ECtHR, 18 December 2012)

Avnish Bajaj v State of Delhi (2005) 116 DLT 427

Balsytė-Lideikienė v Lithuania App no 72596/01 (ECtHR, 4 November 2008)

Case 8611/12 Corte d'Appello di Milano (21 December 2011), Sezione Prima Penale (Italy)

Case C-360/10 *SABAM v Netlog NV* (ECJ, 16 February 2012)

Case C-70/10 *Scarlet Extended SA v Société belge des auteurs, compositeurs et éditeurs SCRL (SABAM)* (ECJ, 24 November 2011)

Cf Bolger v Youngs Drug Products Corp 463 US 60 (1983)

Consolidated Edison Co v Public Service Comm'n of New York 447 US 530, 447 US 542 (1980)

Delfi AS v Estonia App no 64569/09 (ECtHR, 16 June 2015)

Féret v Belgium App no 15615/07 (ECtHR, 16 July 2009)

Frisby v Schultz 487 US 474 (1988)

Garaudy v France App no 65831/01 (ECtHR, 24 June 2003)

Good v Botswana (2010) AHRLR 43 (ACmHPR 2010)

Handyside v UK App no 5493/72 (ECtHR, 7 December 1976)

Jersild v Denmark App no 15890/89 (ECtHR, 23 September 1994)

Joint cases *Google France SARL and Google Inc. v Louis Vuitton Malletier SA* C-236/08, *Google France SARL v Viaticum SA and Luteciel SARL* C-237/08 and *Google France SARL v Centre national de recherche en relations humaines (CNRRH) SARL and Others* C-238/08

Kedar Nath Singh v State of Bihar 1962 Supp (2) SCR 769

Leroy v France App no 36109/03 (ECtHR, 2 October 2008)

Magyar Tartalomszolgáltatók Egyesülete And Index.Hu Zrt v Hungary App no 22947/13 (ECtHR, 2 February 2016)

Mouvement raëlien suisse v Switzerland App no 16354/06 (ECtHR, 13 July 2012)

Norwood v UK App no 23131/03 (ECtHR, 16 November 2004)

Perinçek v Switzerland App no 27510/08 (ECtHR, 15 October 2015)

Pihl v Sweden App no 74742/14 (ECtHR, 7 February 2017)

Pravasi Bhalai Sangathan v. Union of India and Ors (2014) SC 1591

R v Keegstra 3 SCR 697 (1990)

Ramesh v Union of India AIR 1988 SC 775

Ramji Lal v State of Uttar Pradesh 1957 AIR 620

Reno v ACLU 521 US 844 (1997)

Sheppard and Whittle [2010] EWCA Crim 824

Shreya Singhal vs. Union of India AIR 2015 SC 1523

Steel and Morris v the United Kingdom App no 68416/01 (ECtHR, 15 February 2005)

TGI Paris, 20 novembre 2000, *UEJF, LICRA et MRAP (intervenant volontaire) c/ Yahoo! Inc. et Yahoo France*

Vejdeland and others v Sweden App no 1813/07 (ECtHR, 9 February 2012)

Whitney v California 274 US 357 (1927)

International and regional instruments

Additional Protocol to the Convention on Cybercrime, concerning the criminalization of acts of a racist and xenophobic nature committed through computer systems (adopted 28 January 2003, entered into force 01 March 2006) ETS No 189

African Charter on Human and Peoples' Rights (adopted 27 June 1981, entered into force 21 October 1986) (1982) 21 ILM 58 (AfCHR)

American Convention on Human Rights (adopted 22 November 1969, entered into force 18 July 1978) 1144 UNTS 123 (ACHR)

Convention on the Prevention and Punishment of the Crime of Genocide (adopted 9 December 1948, entered into force 12 January 1951)

Council of Europe Committee of Ministers, 'Appendix to Recommendation CM/Rec(2018)2 Guidelines for States on actions to be taken vis-à-vis internet intermediaries with due regard to their roles and responsibilities' (2018)

Council of Europe Committee of Ministers, 'Recommendation to member states on a new notion of media' (2011) CM/Rec (2011)7

Council of Europe Declaration on Freedom of Communication on the Internet (adopted 28 May 2003)

Council of Europe, 'Recommendation No. R (97) 20 of the Committee of Ministers to Member States on Hate Speech' (1997) R (97) 20

Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects on information society services, in particular electronic commerce, in the Internal Market, OJ L 178, 17.7.2000

European Convention on Human Rights (adopted 4 November 1950, entered into force 3 September 1953) (ECHR)

International Convention on the Elimination of All Forms of Racial Discrimination (adopted 21 December 1965, entered into force 4 January 1969) (ICERD)

International Covenant on Civil and Political Rights (adopted 16 December 1966, entered into force 23 March 1976) 999 UNTS 171 (ICCPR)

Joint Declaration on Freedom of Expression and "Fake News", Disinformation and Propaganda by the United Nations Special Rapporteur on Freedom of Opinion and Expression, the Organization for Security and Co-operation in Europe Representative on Freedom of the Media, the Organization of American States Special Rapporteur on Freedom of Expression and the African Commission on Human and Peoples' Rights Special Rapporteur on Freedom of Expression and Access to Information (2017)

Joint Declaration on the Internet and On Anti-Terrorism Measures by the UN Special Rapporteur on Freedom of Opinion and Expression, the OSCE Representative on Freedom of the Media and the OAS Special Rapporteur on Freedom of Expression (2005)

Universal Declaration of Human Rights (adopted 10 December 1948) UNGA Res 217 A(III) (UDHR)

Statutes

Criminal Code in the version promulgated on 13 November 1998, Federal Law Gazette [Bundesgesetzblatt] I p. 3322, last amended by Article 1 of the Law of 24 September 2013, Federal Law Gazette I p. 3671 and with the text of Article 6(18) of the Law of 10 October 2013, Federal Law Gazette I p 3799

Information Technology (Intermediaries Guidelines) Rules, 2011, Gazette of India, pt III sec 4 (11 April 2011)

Information Technology Act 2000, Gazette of India

Netzwerkdurchsetzungsgesetz (NetzDG)

The Indian Penal Code, Gazette of India

Other sources

'13 infamous cases in which Section 66A was misused' (*IndiaToday.in*, March 25, 2015) <https://www.indiatoday.in/india/story/section-66a-cases-how-it-curbed-245739-2015-03-24> accessed 18 October 2018

'EU digital Commissioner attacks German's hate speech bill' *Financial Times* (London) <https://www.ft.com/content/1407bcd8-0d68-11e7-b030-768954394623> accessed 29 June 2018

'Russian bill is copy-and-paste of Germany's hate speech law' *Reporters Without Borders* (Paris, 19 July 2017) <https://rsf.org/en/news/russian-bill-copy-and-paste-germanys-hate-speech-law> accessed 30 October 2018

Answers to the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression in regard to the Act to Improve Enforcement of the Law in Social Networks (Network Enforcement Act), provided by the Federal Government of Germany (2017) <https://www.ohchr.org/Documents/Issues/Opinion/Legislation/GermanyReply9Aug2017.pdf> accessed 2 November 2018

Article 19, 'Germany: Draft Bill on the Improvement of Enforcement of Rights in Social Networks' (2017) <https://www.article19.org/data/files/medialibrary/38723/170426-Germany-Hate-Speech-Law-Draft-Analysis.pdf> accessed 29 June 2018

Article 19, 'Hate speech' Explained – A toolkit' (2015) <https://www.article19.org/data/files/medialibrary/38231/'Hate-Speech'-Explained---A-Toolkit-%282015-Edition%29.pdf> accessed 4 February 2018

Article 19, 'Internet Intermediaries: Dilemma of Liability' (2013) <https://www.article19.org/wp-content/uploads/2018/02/Intermediaries_ENGLISH.pdf> accessed 7 February 2018

Article 19, 'The Camden Principles on Freedom of Expression and Equality' (2009) <https://www.article19.org/data/files/pdfs/standards/the-camden-principles-on-freedom-of-expression-and-equality.pdf> accessed 31 January 2018

Article 19, 'Towards an interpretation of article 20 of the ICCPR: Thresholds for the prohibition of incitement to hatred' (2010) <http://www.ohchr.org/Documents/Issues/Expression/ICCPR/Vienna/CRP7Callamard.pdf> accessed 1 February 2018

Benesch S, 'Dangerous Speech: A Proposal to Prevent Group Violence' (2012) World Policy Institute <http://www.worldpolicy.org/sites/default/files/Dangerous%20Speech%20Guidelines%20Benesch%20January%202012.pdf> accessed 2 February 2018

BIU – Bundesverband Interaktive Unterhaltungssoftware, 'Opinion on the draft bill of the German Federal Ministry of Justice and Consumer Protection regarding an act for improving law enforcement on social networks (Netzwerkdurchsetzungsgesetz, NetzDG)' (2017) <http://ec.europa.eu/growth/tools-databases/tris/en/search/?trisaction=search.detail&year=2017&num=127> accessed 31 October 2018

Borzenko A and Dmitriev D, 'Russian lawmakers drafted a new version of their latest lousy idea to regulate social media. But just how bad is it?' *Meduza* (Riga, 9 April 2018) <https://meduza.io/en/cards/russian-lawmakers-drafted-a-new-version-of-their-latest-lousy-idea-to-regulate-social-media-but-just-how-bad-is-it> accessed 30 October 2018

CDT, 'Shielding the Messengers: Protecting Platforms for Expression and Innovation' (2012) p. 3 <https://www.cdt.org/files/pdfs/CDT-Intermediary-Liability-2012.pdf> accessed 7 February 2018

Chakrabarti G, Dama S, 'Intermediary Liability and Hate Speech' <https://www.law.uw.edu/media/1395/india-intermediary-liability-of-isps-hate-speech.pdf> accessed 15 October 2018

Chawla K, 'Shreya Singhal, and How Intermediaries are Simply Intermediaries Once Again – Striking Down the Chilling Effect' (*Tech Law Forum*, 30 March 2015) <https://techlawforum.wordpress.com/2015/03/30/shreya-singhal-and-how-intermediaries-are-simply-intermediaries-once-again-striking-down-the-chilling-effect/> accessed 13 October 2018

Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, 'Tackling Illegal Content Online: Towards an enhanced responsibility of online platforms' COM (2017) 555

Dara R, 'Intermediary Liability in India: Chilling Effects on Free Expression on the Internet' (*The Centre for Internet & Society*, 27 April 2012) <https://cis-india.org/internet-governance/chilling-effects-on-free-expression-on-internet> accessed 13 October 2018

Dara R, 'Intermediary Liability in India: Chilling Effects on Free Expression on the Internet' (2011) <https://cis-india.org/internet-governance/intermediary-liability-in-india.pdf> accessed 13 October 2018

Eddy M, Scott M, 'Delete Hate Speech or Pay Up, Germany Tells Social Media Companies' *New York Times* (New York, 30 June 2017) <https://www.nytimes.com/2017/06/30/business/germany-facebook-google-twitter.html> accessed 4 November 2018

EDRi, 'Recommendations on the German bill "Improving Law Enforcement on Social Networks" (NetzDG)' (2017) <https://edri.org/files/consultations/tris_netzdg_edricontribution_20170620.pdf> accessed 1 November

EU Commission, 'Code of Conduct on countering illegal hate speech online: First results on implementation' (2016) <http://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupDetailDoc&id=29738&no=1> accessed 12 November 2018

European Commission, 'Liability of Internet Intermediaries in Legal analysis of a Single Market for the Information Society (SMART 2007/0037)' (2011) <https://ec.europa.eu/digital-single-market/news/legal-analysis-single-market-information-society-smart-20070037> accessed 12 November 2018

Facebook, 'Community standards: Helping to keep you safe' <https://www.facebook.com/communitystandards#dangerous-organizations> accessed 1 February 2018

Facebook, 'Statement of Rights and Responsibilities' <https://www.facebook.com/legal/terms> accessed 1 February 2018

Garner A. B (ed.), *Black's Law Dictionary* (8rd edn, The West Group 2014)

Ginkel van B, 'Incitement to Terrorism: A Matter of Prevention or Repression?' (2011) ICCT Research Paper <http://www.icct.nl/download/file/ICCT-Van-Ginkel-Incitement-To-Terrorism-August-2011.pdf> accessed 2 February 2018

Hatebase, 'Most common hate speech' <https://www.hatebase.org/popular> accessed 1 February 2018

Index on Censorship, 'Index supports referral request in Delfi v. Estonia' (*Xindex*, 14 January 2014) <http://www.indexoncensorship.org/2014/01/index-supports-referral-request-delfi-v-estonia/> accessed 3 October 2018

Interview with Drew Boyd, Director of Operations, The Sentinel Project for Genocide Prevention (24 October 2014)

Joshi D, 'Indian Intermediary Liability Regime Compliance with the Manila Principles on Intermediary Liability' <https://cis-india.org/internet-governance/files/indian-intermediary-liability-regime> accessed 21 October 2018

Llanso E, 'German Proposal Threatens Censorship on Wide Array of Online Services' *CDT* (7 April 2017) <https://cdt.org/blog/german-proposal-threatens-censorship-on-wide-array-of-online-services/> accessed 29 June 2018

Manila Principles on Intermediary Liability: Best Practices Guidelines for Limiting Intermediary Liability for Content to Promote Freedom of Expression and Innovation (2015), <https://www.manilaprinciples.org> accessed 10 October 2018

McGonagle T, 'The Council of Europe against online hate speech: Conundrums and challenges' <https://rm.coe.int/16800c170f> accessed 13 October 2017

Ministry of Communications and Information Technology Department of Electronics & Information Technology, 'Clarification on The Information Technology (Intermediary Guidelines) Rules, 2011 under section 79 of the Information Technology Act, 2000' (2013) <http://cyberlawindia.com/wp-content/uploads/2014/06/Clarification-79rules1.pdf> accessed 16 October 2018

Mumo M, 'Social media sites to delete hate mongers' accounts in a day' *Daily Nation* (15 July 2017) <https://www.nation.co.ke/business/Social-media-sites-to-delete-hate-mongers-accounts-in-a-day/996-4016026-qo0k2n/index.html> accessed 30 October 2018

OECD, 'The Economic and Social Role of Internet Intermediaries' (2010) <https://www.oecd.org/internet/ieconomy/44949023.pdf> accessed 7 February 2018

Office of the Special Rapporteur for Freedom of Expression of the Inter-American Commission on Human Rights, 'Freedom of Expression and the Internet' (2013)

Opinion of David Kaye Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression regarding the German Draft Law Netzdurchführungsgeset <https://www.ohchr.org/Documents/Issues/Opinion/Legislation/OL-DEU-1-2017.pdf> accessed 29 June 2018

OSCE Representative on Freedom of the Media, 'Amsterdam Recommendations on the Freedom of the Media and the Internet' from 14 June 2003 available at <https://www.osce.org/fom/13854?download=true> accessed on 26 September 2018

OSCE, 'Legal Review of the Draft Law on Better Law Enforcement in Social Networks' (2017) <https://www.osce.org/fom/333541?download=true> accessed 29 June 2018

Packer G, 'Mute Button' *The New Yorker* (New York, 13 April 2015) <https://www.newyorker.com/magazine/2015/04/13/mute-button> accessed 7 November 2018

Panday J, 'The Supreme Court Judgment in Shreya Singhal and What It Does for Intermediary Liability in India?' (*The Centre for Internet & Society*, 11 April 2015) <https://cis-india.org/internet-governance/blog/sc-judgment-in-shreya-singhal-what-it-means-for-intermediary-liability> accessed 15 October 2018

Ribeiro J, 'Experts criticize arrest of Baazee.com CEO' *Macworld* (20 December 2004) <https://www.macworld.com/article/1041530/experts.html> accessed 16 October 2018

Rohleder B, 'Germany set out to delete hate speech online. Instead, it made things worse' *The Washington Post* (Berlin, 20 February 2018) <https://www.washingtonpost.com/news/theworldpost/wp/2018/02/20/netzdg/?utm_term=.7a50002de4bf> accessed 2 November 2018

Sajo A, 'The Fate of Human Rights in Indifferent Societies' (Demise of Constitutionalism Conference, Budapest, May 2018)

Scolaro M C and Morganteen J, 'Facebook blocks more content here than in any other country' (CNBC, 13 November 2015) <https://www.cnbc.com/2015/11/13/facebook-blocks-more-content-here-than-any-other-country.html> accessed 17 October 2018

The Guardian and LSE, 'Reading the Riots Investigating England's Summer of Disorder' (2011) <http://eprints.lse.ac.uk/46297/1/Reading%20the%20riots(published).pdf> accessed 18 October 2018

Thelle H M and others, 'Closing the Gap – Indian Online Intermediaries and a Liability System Not Yet Fit for Purpose' (Global Network Initiative, 2014) <https://www.copenhageneconomics.com/dyn/resources/Publication/publicationPDF/1/251/0/Closing%20the%20Gap%20-%20Copenhagen%20Economics_March%202014.pdf> accessed 12 October 2018

Twitter, 'Hateful conduct policy' <https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy> accessed 1 February 2018

Ulvik M and Pavli D, 'Case Watch: A Strasbourg Setback for Freedom of Expression in Europe' (*Open Society Foundations*, 22 October 2013) <https://www.opensocietyfoundations.org/voices/case-watch-strasbourg-setback-freedom-expression-europe> accessed 11 October 2018

UNESCO, 'Comments on the Draft for an Act improving Law Enforcement on Social Networks (NetzDG)' (2017) <https://www.hans-bredow-institut.de/uploads/media/default/cms/media/4c70991cc1654caa2b47b509bad7bd1328824391.pdf> accessed 31 October 2018

UNESCO, 'Countering online hate speech' (2015) <http://unesdoc.unesco.org/images/0023/002332/233231e.pdf> accessed 25 November 2018

UNESCO, 'Fostering Freedom Online: The role of Internet Intermediaries' (2014) <http://unesdoc.unesco.org/images/0023/002311/231162e.pdf> accessed 7 February 2018

UNESCO, 'World Trends in Freedom of Expression and Media Development' (2015) <http://unesdoc.unesco.org/images/0023/002349/234933e.pdf> accessed 9 November 2018

UNGA 'Rabat Plan of Action on the prohibition of advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence' (2012) A/HRC/22/17/Add.4

UNGA, 'Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression' (2012) UN Doc A/67/357

UNGA, 'Report of the Special Rapporteur to the General Assembly on hate speech and incitement to hatred' (2012) UN Doc A/67/357

UNGA, 'The promotion, protection and enjoyment of human rights on the Internet' (2012) A/HRC/RES/20/8

UNHRC, 'General Comment No. 11: Prohibition of propaganda for war and inciting national, racial or religious hatred (Art. 20)' (1983)

UNHRC, 'Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression Frank La Rue' (2011) UN Doc A/HRC/17/27

Voorhoof D, 'Delfi AS v. Estonia: Grand Chamber confirms liability of online news portal for offensive comments posted by its readers' (*Strasbourg Observers*, 18 June 2015)

<https://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 10 October 2018

Voorhoof D, '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) <https://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 11 October 2018

Voorhoof D, Lievens E, 'ECtHR confirms and tempers Delfi judgment: operators of Internet portals not liable for dissemination of offending - but not "clearly unlawful" - user comments' (*ECHR BLOG*, 15 February 2016) <http://echrblog.blogspot.com/2016/02/offensive-online-comments-new-ecthr.html> accessed 11 October 2018

YouTube, 'Hate speech' <https://support.google.com/youtube/answer/2801939?hl=en> accessed 1 February 2018