

IS SOCIAL MEDIA HARNESSING HATE SPEECH: WHETHER A CRIME OR FREEDOM OF EXPRESSION?

Namita Bhardwaj

Senior Research Fellow,

Department of Laws, Panjab University, Chandigarh

Abstract

Social Media platforms have developed to be the most important source of information dissemination. Users of social media can upload pictures, videos, etc on these platforms. It has brought the world together virtually. It allows people to express themselves easily, effectively and across geographical boundaries. However, at the same time, social media is being used as a launchpad by hate mongers to spread hatred for another section of society. The difference between one's ability to express criticism about some situation from one's ability to incite hate through such criticism is very bleak. In the present paper, the researcher has tried to address this issue by dividing the research into five parts. In the first part, the researcher has explained hate speech. In the second part, the international stand of various countries in dealing with hate speech is discussed. In the third part, the efforts of social media platforms in curbing hate speech is discussed. In the fourth part, the need for regulating hate speech is discussed. The fifth part contains of the conclusion and suggestions.

Keywords: Hate Speech; Social Media; Internet; Hate Crime; Freedom of speech and expression

INTRODUCTION

Internet and social media have been one of the biggest crusaders of twenty first century in taking forward the individual's right of freedom of expression by providing convenient platforms where an individual can express his or her feeling regarding any issue, with the assurance that the internet technology would disseminate the information all over the world at the click of a button. However, this boon has in recent past shown disturbing trends where internet and social media platforms are used by hate mongers to preach hatred and incite people against other sections of society. In 2018, the Pittsburgh synagogue attack in USA which led to the death of.; the attacker was found to be an active user of a banned platform "Gab" where people shared anti- sementic views.¹ The horrible 2019 New Zealand Mosque attack which led to the killing of 50 people was live- streamed by the attacker on facebook and he also distributed a hate- filled manifesto through the same platform.² In Myanmar, online hate speech against Rohingyas led to ...; while in India hate speech resulted in mob- lynching. In light of the rising hate speech incited hate crime and violence incidents and as hate speech is positioned at the crossroad of conflicting principles of freedom of expression at one hand and right to protection against discrimination at other hand; it is imperative to cautiously define hate speech. The Cambridge Dictionary defines hate speech as, "public speech that expresses hate or encourages violence towards a person or group based on something such as race, religion, sex or sexual orientation"³ The European Commission against Racism and Intolerance (ECRI) has defined it as, "Hate speech covers many forms of expressions which spread, incite, promote or justify hatred, violence and discrimination against a person or group for a variety of reasons."⁴ The Law Commission of India in its 267th report has defined hate speech as "an expression which is likely to cause distress or offend other individuals on the basis of their association with a particular group or incite hospitality towards them."⁵ Thus, hate speech can be construed any expression which is meant to incite others against another section of society in order to discriminate or dehumanize that section can be construed to be hate speech. Criticism of others without any intent to incite would not qualify to be hate speech.

International stand regarding hate speech: Adverse impact of hate speech on human rights has been recognized in international documents and treaties. The first important international document to recognize hate speech is Universal Declaration of Human Rights (hereinafter referred UDHR) in 1948. The magna- carta of human rights in Article 7 has implicitly extended protection from hate speech by providing right to equal

¹ Jonathan Peters, "How the law protects hate speech on social media" Columbia Journalism Review (CJR) Nov. 2, 2018. Available at: <https://www.cjr.org/analysis/gab-hate-speech.php>

² Damien Cave, "Australia Passes Law to Punish Social Media Companies for Violent Posts". The New York Times. On April 3, 2019. Available at: <https://www.nytimes.com/2019/04/03/world/australia/social-media-law.html>. Accessed on: 1st October, 2019

³ Definition of Hate Speech. Available at: <https://dictionary.cambridge.org/dictionary/english/hate-speech>. (Last viewed on 25th October, 2019)

⁴ Hate Speech and Violence. Available at: <https://www.coe.int/en/web/european-commission-against-racism-and-intolerance/hate-speech-and-violence>. (Last viewed on 25th October, 2019)

⁵ Law Commission of India, 267th Report on Hate Speech (March, 2017)

protection and protection against any discrimination.⁶ On the basis of UDHR, 'The International Covenant on Civil and Political Rights' (hereinafter referred ICCPR) has extended vigorous protection against hate speech. Article 20 of ICCPR provides for prohibition of advocacy to hatred leading to incitement to discrimination, hostility or violence.⁷ While, Article 19 of ICCPR provide for freedom of expression to everyone; it is not absolute and is subject to reputation of others and conditions like national security, public order, etc.⁸ Although both the provisions are conflicting, it has to be interpreted to suggest that for exercising right to expression under Article 19, it should not violate Article 19(3) and Article 20 of ICCPR. Other international instruments like the International Convention on the Elimination of All Forms of Racial Discrimination (hereinafter referred as ICERD), The Convention on the Prevention and Punishment of the Crime of Genocide (hereinafter referred as Genocide Convention) and the Convention on the Elimination of All Forms of Discrimination against Women (hereinafter referred as CEDAW) provide for protection against hate speech targeted to incite committing of genocide or targeting racial discrimination or speech targeting women.

The above- mentioned international instruments act as the basis for the states to form its own legislations.⁹ In United Kingdom, hate speech is regulated under varied laws.¹⁰ The Crime and Disorder Act, 1998 punishes aggravated offences arising from racial or religious grounds like threatening or abusive and insulting conduct against any race or religion.¹¹ The Public Order Act, 1986 punishes conduct that is intended to create hatred on grounds of race, religion or sexual orientation.¹² Part 1 of the Malicious Communications Act, 1988 punishes sending indecent and grossly offensive communications which are made with the intention of causing distress in the society.¹³ Section 127 of the Communications Act, 2003 prohibits sending grossly offensive or obscene messages on public electronic communications network. Despite of so many laws, there is no legislation currently addressing the issue of online hate speech in U.K.. However, the report on Hate Crime released in April 2017, "Hate Crime: abuse, hate and extremism online" made the following recommendations. Firstly, it recommended that the law regarding online hate speech should be updated. Secondly, the social media companies must review their standards for dealing with spread of hate speech from their platforms.¹⁴

In European Union, The 2016 European Union Code on countering illegal hate speech online involves agreement between European Union and four major IT companies, Facebook, Twitter, Youtube and Microsoft to set a target for removal of hate speech within 24 hours of its notification.¹⁵ In 2017, the Commission reported progress in the take- downs of hate speech which witnessed an increase from 28% to 59% in a 6 month period.¹⁶

In USA, there is no proper protection against hate speech as the first amendment to U.S. constitution supports freedom of speech. The first amendment protection prohibits the government from making laws that abridges freedom of speech and expression. However, in addressing hate speech, two landmark judgments of the U.S. Supreme Court are relevant in this context. The U.S. Supreme Court in *Chaplinsky v. New Hampshire*¹⁷ differentiated between classes of speech and held that certain expressions like "fighting words, obscenities, certain profane and slanderous speech" are outside the purview of protection of first amendment. Following this, the U.S. Supreme Court in *Brandenburg v. Ohio*¹⁸ laid down the "present danger test" for ascertaining whether a speech qualifies to be hate speech. It lays down that where the speech directly advocates inciting or producing imminent lawless action, then such an action should not get the protection of freedom of speech and expression.¹⁹ Thus, hate speech regulation in USA is mainly self- regulated by the internet industry and on the legal front, the aggrieved can seek remedies under tort law²⁰ and criminal law²¹.

⁶ "All are equal before the law and are entitled without any discrimination to equal protection of the law. All are entitled to equal protection against any discrimination in violation of this Declaration and against any incitement to such discrimination". Article 7, UDHR

⁷ Article 20, ICCPR. Article 20(2) states, "Any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law/"

⁸ Article 19(3), ICCPR

⁹ Iginio Gagliardone, et. al., "Countering Online Hate Speech" UNESCO 2015. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000233231> (Last viewed on 20th October, 2019)

¹⁰ Catherine O' Regan, "Hate Speech Online: an (Intractable) Contemporary Challenge?" Current Legal Problems, Vol. 71, No. 1 (2018), pp. 403-429. Available at: doi:10.1093/clp/cuy012

¹¹ Ibid @pg 419

¹² Ibid

¹³ Supra 10

¹⁴ Ibid @pg 421

¹⁵ "Code of Conduct on countering illegal hate speech online". Available at: https://ec.europa.eu/info/sites/info/files/code_of_conduct_on_countering_illegal_hate_speech_online_en.pdf

¹⁶ Lilian Edwards, "Law, policy and Internet" (HART, Oxford, 2018)

¹⁷ 315 U.S. 568 (1942)

¹⁸ 395 U.S. 44 (1969)

¹⁹ Supra 5

²⁰ US DCA can be used by the aggrieved for punishing the accused for defamation.

Australia passed a stringent law regarding online hate speech regulation “Criminal Code Amendment (sharing of abhorrent violent material) Bill” on March, 2019 which imposes huge fines on social media companies and jail for the company executives in the event of failure by the company to remove from their platforms “abhorrent abusive material”. The legislation was passed hastily in the backdrop of 2019 New Zealand mass-killings by an Australian white nationalist.

In India, efforts by the legal fraternity for curbing hate speech begun after the *Pravasi Bhalai Sangathan v. Union of India & Ors.*²² in which the honorable supreme court proposed a report from Law commission report on the same. Consequently, the law commission in its 276th report published in March, 2017 while explaining hate speech, recognized internet as a mode of transmission of hate speech. It states, “...”. In the recent landmark judgment of *Shreya Singhal v Union of India*, the apex court has differentiated between three forms of speech, ie, discussion, advocacy and incitement. The court in this case held that a speech can be limited on grounds of exception mentioned in Article 19(2) only when it reaches the threshold of incitement. All other forms of speech, even if offensive or unpopular have to be protected under Article 19(1)(a). The court has said, “Incitement is the key to determining the constitutionality of restriction of free speech.”²³ Recently, the Indian Government has framed new social media rules which empower the government to order platforms to take down posts within twenty-four hours based on a wide range of offenses, as well as to obtain the identity of the user.²⁴

Efforts of internet platforms in curbing hate speech- The internet and social media platforms are tackling the problem of hate speech through following the national laws as discussed above and through self regulation. Most of these platforms have defined what instances amount to hate speech and have also defined the measures undertaken for responding to it. In case a user posts content that amounts to hate speech, such content may be removed or given a restricted access. After the 2019 New Zealand mass killings, Facebook has added new restrictions for live video streaming by detecting and removing worrying images and videos from the internet.²⁵ YouTube also prohibits hate speech on its platform which has the potential of attacking or demeaning a group based on race or ethnicity, nationality, sexual orientation, sex, gender identity or serious disabilities or diseases.²⁶ However, the problem with self-regulation is that editorial authority given to platforms in absence of a regulatory framework may cause infringement of freedom of speech and expression.

Regulation of online hate speech-

Article 19(2) of ICCPR states that freedom of expression is available for information of all kinds, whether oral, written or in print, or in form of art or through any other media of his choice.²⁷ This provision implicitly implies that ICCPR applies to new forms of technology, particularly internet.²⁸ Thus, online speech posted on social media platforms have to adhere to principles laid down in ICCPR of freedom of speech and expression under Article 19 and prohibition of advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence under Article 20 of ICCPR. The United Nations High Commissioner for Human Rights in 2012 recommended the adoption of “Rabat Plan of Action on the prohibition of advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence” (hereinafter referred as Rabat plan of action) for regulating restrictions which maybe placed on speech in light of Article 19 and Article 20 of ICCPR.²⁹ For determining hate speech, a six- part threshold test has been laid down.:

1. Context- Analysis of the context should place the speech act within the social and political context prevalent at the time the speech was made and disseminated.³⁰

²¹ In case of any consequent hate crime emerging from hate speech, the US penal law can be used for prosecuting the accused.

²² AIR 2014 SC 1591

²³ Supra 5

²⁴ “Hate Speech on Social Media: Global Comparisons” by Zachary Laub. Link-

<https://www.cfr.org/backgrounder/hate-speech-social-media-global-comparisons>

²⁵ Cade Metz and Adam Satariano, “Facebook Restricts Live Streaming After New Zealand Shooting” *The New York Times*, May 14, 2019. Available at: <https://www.nytimes.com/2019/05/14/technology/facebook-live-violent-content.html>. (Last viewed on 25th October, 2019)

²⁶ Supra 8. @ Pg 29

²⁷ Article 19(2), ICCPR.

²⁸ Supra 8

²⁹ @PG 11. https://www.ohchr.org/Documents/Issues/Opinion/SeminarRabat/Rabat_draft_outcome.pdf

³⁰ Annual Report of the United Nations High Commissioner for Human Rights on the expert workshops on the prohibition of incitement to national, racial or religious hatred. Available at: https://www.ohchr.org/Documents/Issues/Opinion/SeminarRabat/Rabat_draft_outcome.pdf. (Last viewed on 25th October, 2019)

2. Speaker- The speaker's position or status in the society should be considered, specifically the individual's or organization's standing in the context of the audience to whom the speech is directed.³¹
3. Intent- Article 20 of the International Covenant on Civil and Political Rights anticipates intent. Negligence and recklessness are not sufficient for an act to be an offence under article 20 of the Covenant, as this article provides for "advocacy" and "incitement" rather than the mere distribution or circulation of material. In this regard, it requires the activation of a triangular relationship between the object and subject of the speech act as well as the audience.³²
4. Content and form- The content of the speech constitutes one of the key foci of the court's deliberations and is a critical element of incitement. Content analysis may include the degree to which the speech was provocative and direct, as well as the form, style, nature of arguments deployed in the speech or the balance struck between arguments deployed.³³
5. Extent of the speech act- Extent includes such elements as the reach of the speech act, its public nature, its magnitude and size of its audience. Other elements to consider include whether the speech is public, what means of dissemination are used, for example by a single leaflet or broadcast in the mainstream media or via the Internet, the frequency, the quantity and the extent of the communications, whether the audience had the means to act on the incitement, whether the statement (or work) is circulated in a restricted environment or widely accessible to the general public.³⁴
6. Likelihood, including imminence- Incitement, by definition, is an inchoate crime. The action advocated through incitement speech does not have to be committed for said speech to amount to a crime. Nevertheless, some degree of risk of harm must be identified. It means that the courts will have to determine that there was a reasonable probability that the speech would succeed in inciting actual action against the target group, recognizing that such causation should be rather direct.³⁵

The Rabat plan of Action provides guidelines for states and not social media platforms. Hence, they are not duty bound to follow the recommendations of the plan of action. Another important document in this regard is the "The UN Guiding Principles on Business and Human Rights" framed by United Nations Office of the High Commissioner. It emphasizes corporate responsibility in upholding human rights. In the event of violation of human rights, these guidelines are to be referred to by internet companies. Recently, the report of United Nations "Special Rapporteur on the promotion and protection of the right to freedom of speech and expression"³⁶ has made the recommendations addressing both the states and companies in addressing online hate speech. The important recommendation made for states is that regulation of online content should be smart regulation and not heavy-handed regulation. It states, "States should only seek to restrict content pursuant to an order by an independent and impartial judicial authority, and in accordance with due process and standards of legality, necessity and legitimacy. States should refrain from imposing disproportionate sanctions, whether heavy fines or imprisonment, on Internet intermediaries, given their significant chilling effect on freedom of expression."³⁷ Recognising the significance of internet companies in maintaining freedom of speech and expression, the report recommends an authoritative global standard to be followed. It suggests that the terms and policies of the companies should be framed in the backdrop of human rights law. The internet companies should also be made more accountable by imposing minimum levels of consistency, transparency and accountability to commercial content moderation.³⁸

CONCLUSION AND SUGGESTIONS

The increase in hate crime caused due to the rapid dissemination of hate speech through internet and social media platforms has made it a concern of immediate attention by the legal fraternity. Different countries have addressed the issue of hate speech with different perspective. In USA, the approach is more towards protection of freedom of speech and expression. Whereas in countries like Australia, strict laws have been passed to

³¹ Ibid

³² Ibid

³³ Ibid

³⁴ Ibid

³⁵ Ibid

³⁶ The report of United Nations "Special Rapporteur on the promotion and protection of the right to freedom of speech and expression. Available at: <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G18/096/72/PDF/G1809672.pdf?OpenElement> (Last viewed on 20th October, 2019)

1.1. ³⁷ "2018 thematic report to the Human Rights Council on content regulation". Available at: <https://www.ohchr.org/EN/Issues/FreedomOpinion/Pages/ContentRegulation.aspx> (Last viewed on 20th October, 2019)

³⁸ Ibid



counter hate speech. A balanced approach in addressing hate speech is required in order to assure proper implementation of human rights as enshrined in UDHR and ICCPR. Thus, a global regulation on hate speech establishing rules for states as well as internet and social media companies for assuring freedom of speech and expression without causing any discrimination or hatred against a particular section of society.