
SOCIAL MEDIA AND UNREFINED SPEECH: ANALYZING BOUNDARIES IN THE CONTEMPORARY DIGITAL WORLD

Sudiksha Kejriwal, Research Scholar, Faculty of Law, Banaras Hindu University

ABSTRACT

There is considerable perplexity in identifying content as hateful, as hate speech is not defined uniformly and the determinations of speech being offensive or as to what actually conveys hate are vague and subjective. We do not have a definite criterion or set boundary to demarcate between hateful and non-hateful speech. A significant arena for dissemination of hate speech today is social media, where it frequently appears on platforms like Facebook, X (formerly Twitter), Instagram, YouTube, etc. This represents a pressing problem for online forums that allow user-generated content, and current approaches for handling such inadvertent behavior are still inefficient. The paper aims to review the theoretical frameworks in the literature on hate speech and trace out its development over the years. The risks associated with hate speech cannot be fully understood by solely focusing on the content of hateful speech acts; rather, the broader communication dynamics and relationships behind these acts also needs to be carefully considered. Several studies strive to develop effective tools for detection; however, significant challenges remain in identifying harmful online behavior, particularly in addressing issues of anonymity and going beyond merely detecting offensive words. The way language has been transformed by the evolution of the internet has turned finding the definitions between hate speech and exercising free speech into one of the most difficult things ever to grapple.

Keywords: Hate speech, regulation, offensive, free speech, and social media.

"Our freedom of speech is being eroded and mauled through twisting and turning the law if not abusing it altogether."

-Justice Madan B. Lokur

INTRODUCTION

The recent advances in technology have made society increasingly dependent on it and will even surge in the future. Amongst these technologies, the Internet poses major challenges to the individual security of people; it has become much easier for hateful individuals to find others who share their views, resulting in a rise in societal hate. The unchecked proliferation of hate poses a significant threat to society, disproportionately impacting marginalized individuals and communities. For our society to become peaceful and safe, hate can no longer be tolerated and needs to be eliminated immediately.

Hate speech, being a global concern, needs to be understood from multiple perspectives. Our civilization regrettably includes hate speech, which has now found a new platform in the modern era of real-time digital culture. It is apparent today that social media such as Facebook, X, YouTube and others have become the new playground for regular exchanges tinged with unpleasant hatred. We can say that prevention of hate speech will require an exhaustive study and analysis of hate speech.

Hate, as we consider to be more than an emotion, or state of mind is also known to serve as an evolutionary function and we should not view it as a solely evil entity. This is an angry or resentful feeling attitude towards certain people or things¹. Hatred can last seconds or continue to build for years, strong or mild. We have gone through recent development of social media over the recent couple of years that has actually changed the whole face of communications. Social media offer a low-cost communication medium that connects billions of people worldwide, allowing people to share their opinions freely, the ability to voice what's on their mind, be it a friend's or foe's. This represents a revolutionary change in our culture. It also serves as a platform for individuals, regardless of age or profession, to express their opinions.

However, while there are numerous advantages, there exist notable disadvantages, social media platforms present a significant challenge, as it amplifies discourses that are detrimental to

¹ Reber, A.S. and Reber, E., "The Penguin dictionary of psychology", New York: Penguin Books (2002)

specific demographic groups. This can occur in various forms, such as hate speech, offensive content, and violence. In recent years, it can be seen that, many countries are now recognizing that hate speech is a serious issue, particularly in light of the difficulty of establishing online barriers to prevent the spread of hate among individuals and communities. This paper aims to explore the prevalence of hate speech online social media, types of hate speech which are most common and widespread, and how does anonymity affect hate speech.

With regards to the social media, all kinds of posts such as those about politics, economy, art, literature, and religion can be shared directly without vast restrictions. Consequently, many posts impact others beyond just the realm of free speech. The records in regional languages are increasing, while causing a number of issues. Controlling such hate speech is vital in the current situation. The study of hate speech has received more attention in recent years, with several articles focusing on its detection on social networking platforms. Because of these platforms, so many issues arise from the instantaneous sharing of people's opinions. Such media promote freedom of speech, but have also opened a dangerous avenue for hate speech.

We have progressed from a time when people shared their ideas mostly orally and within small groups to a time when individuals can freely use a variety of networks to communicate instantly with people across large distances. Additionally, an increasing number of people use online platforms for news sharing as well as social interaction. Expressing one's opinion has now become a natural activity today, and everyone is entitled to do so, unless done inappropriately. Global communication has become more robust and expansive with technological advances. People can now express their opinions through user-generated content on websites, blogs, images, and other platforms via the internet. This new medium of virtual communication enables greater freedom in self-expression as it allows individuals to write, without obligation to reveal their identities. However, like everything, this too has its pros and cons; the dark side is that social media has become a flourishing platform for heated discussions, often accompanied by derogatory and offensive language. Hate speech is not just harsh words; it can occur both offline and online. It can be conveyed through written language or visual representation using symbols, images, memes, emojis, or videos. For example, memes may consist of humorous or seemingly innocent images that upon closer scrutiny, may carry racist or anti-Semitic undertones. The online hate, trolling, cyberbullying and all forms of violence using new technologies, seem to be new phenomena. Our mental resources are limited and must be used cautiously. It has been established that negative emotions often take precedence

in our mental systems, so one must be careful about the attention given to them, ensuring they are not for too long.

In a general sense, hatred is a communication which humiliates an individual, because of their own membership in a certain social group, whether it is an expression, an image, a play, a song, and even speech. It is defined broadly as speech that insults people because of their membership in certain groups, like race, national origin, religion, sex, sexual orientation, or some other distinguishing feature in which members of this group are arbitrarily defined. Several conceptualizations of hate speech exist across countries which makes it difficult to distinguish hate speech from the right of free speech because of no uniform definition. However, an attempt has been made, to derive the idea of hate speech from dictionaries, etc.

Sigmund Freud said that hate is an ego state which wishes to annihilate the source of its own unhappiness, pointing to the issue of self-preservation². Hate speech is defined as a form of public speech which expresses hate or encourages violence against individuals based on race, religion, sex, sexual orientation etc.³. Furthermore, it has been defined as, “speech that has nothing to express except hatred towards somebody’s group; like race in particular’ when communicated in a manner which incites hate and/or incites violence.”⁴

Some online platforms maintains that hate speech is a direct attack against people for their protected characteristics, which refer to race, ethnicity, national origin, religious identity, caste, sexual orientation, sex, gender identity, and disability.⁵ It is defined as the content that promotes violence or hatred targeted at the individual or groups based on certain attributes, including race or ethnicity, national origin, religion, sex, age, veteran status, and sexual orientation/sex, age, veteran status sexual orientation, gender identity etc.⁶ Hate speech is best defined by the speech against a specific social or cultural group that shares characteristics,⁷ i.e. a language which demeans, insults, threatens or targets a person through their identity along with other qualities like sex, sexuality or disability etc.⁸ Hate speech “includes statements

² S Freud, ‘The Instincts and Their Vicissitudes’ Collected Papers, IV London Hogarth Press, p. 135-37 (1915).

³ Cambridge Dictionary, available at <https://www.dictionary.com/>, accessed on 18.08.2024.

⁴ Black’s Law Dictionary, 9th edition, 2009.

⁵ Meta Transparency Centre, available at <https://transparency.meta.com/en-gb/policies/community-standards/hate-speech/>, accessed on 24.06.2024.

⁶ Youtube Community Guidelines, available at <https://transparencyreport.google.com/youtube-policy/featured-policies/hate-speech?hl=en>, accessed on 24.06.2024

⁷ Ivan Hare & James Weinstein, “Extreme Speech and Democracy”, 4 Journal of Media Law, 2009

⁸ The Bureau of Police Research and Development Manual on Cyber Harassment.

intended to incite animosity or hatred for a particular sex, race, color, religion, etc.”⁹ It can be defined as speech that expresses hatred towards a particular person or group.¹⁰ In government documents, hate speech is stated as an incitement of the mass hatred. It is defined as the public incitement to hatred against any person or group of persons on account of race, ethnicity, gender, sexual orientation, religious belief and the like.”¹¹ The European Union in its Legislation¹² has partly defined hate speech as: “public incitement to violence or hatred against a group of persons in general or a member of such a group defined on the basis of race, color, descent, religion or belief, or national or ethnic origin.” It is worth noting that The Council of Europe (Committee of Minister) has interpreted a broader approach of hate speech as covering 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.”¹³ Therefore, it can be said that hate speech is generally an insult towards an individual.¹⁴

By comparing these definitions from various sources, we can see that hate speech aims to attack or diminish specific targets to incite violence or hatred. It conveys no meaning other than the expression of hatred towards a group of people, particularly when there is a possibility that communication may promote violence against individuals characterized by their gender, religion, race, ethnicity, national origin, or sexual orientation. The use of hate speech is often described as morally wrong by society. And as discussed, hate speech laws usually cover not only spoken words, but also words or images printed, published, or posted on the internet.¹⁵ This is particularly crucial today, as anything produced or posted online can reach large audiences. It is essential to analyze how words used can significantly impact individual and society at large, disturbing social harmony and order. Hate speech also affects them in

⁹ John T Nockleby, “Hate Speech” in Encyclopedia of the American Constitution, ed. Leonard W. Levy and Kenneth L. Karst, vol. 3. (2nd ed.), 2000, Detroit: Macmillan Reference US, 1277. Cited in “Library 2.0 and the Problem of Hate Speech,” by Margaret Brown-Sica and Jeffrey Beall, Electronic Journal of Academic and Special Librarianship, Vol. 9 no. 2 (Summer 2008).

¹⁰ Merriam Webster’s Dictionary of English Usage, Springfield, Mass: 1994 retrieved from <https://www.merriam-webster.com/dictionary/hate%20speech>, accessed on 21.03.2024

¹¹ Law Commission of India, *Hate Speech in India*, (Law Com No 267, 2017) para 6.31

¹² European Commission, *Framework Decision on combating certain forms and expressions of racism and xenophobia by means of criminal law*, (2008)

¹³ Council of Europe, *Recommendation No. R 97 (20) of the Committee of Ministers to Member States on Hate Speech*, (30 October 1997)

¹⁴ Gautam Bhatia, “Offend, Shock or Disturb: Free Speech Under the Indian Constitution” 139, (Oxford University Press), Delhi, 1st Edition, 2016.

¹⁵ Jeremy Waldron, “The Harm in Hate Speech”, 37 (Harvard University Press, Cambridge, 2012)

intangible ways, which causes what is known as ‘chilling effect’ to right of free speech and expression, resulting in exclusion from participation in the democratic process and public discourse.¹⁶

CONSTITUTIONAL FRAMEWORK OF SPEECH

“Give me the liberty to know, to utter, and to argue freely according to conscience, above all liberties”.

- John Milton

India is the largest democratic country, and the freedom of speech is its lifeblood like in any other country. The freedom of press and expression serves as a backbone of a functional democracy. The faith in democracy rests upon the old dictums, “Let the people have the truth and freedom to discuss it, and all will go well,”¹⁷ and “No one ought to be molested on account of his opinions...”¹⁸

All citizens shall have the right to freedom of speech and expression.¹⁹ It is recognized as one of the most fundamental requirements for a free democratic nation. This right is borrowed from American Constitution. The right to freedom of speech and expression is also recognized in the Universal Declaration of Human Rights (hereinafter referred to as UDHR) as a human right. This is one of the key rights, which enables the person to achieve self-realization and have freedom with all his might.²⁰

“Your liberty to swing your fist ends just where my nose begins”

Oliver Wendell Holmes

This means you can exercise your rights until they do not harm right of another person. Therefore, it can be seen that this right comes with certain reasonable restrictions as well. In general, freedom of speech and expression refers to the notion that everyone has an inherent

¹⁶Parliamentary Assembly, Council of Europe, *Recommendation 1805 Blasphemy, religious insults and hate speech against persons on grounds of their religion*, (2007)

¹⁷James Bryce, “The Definition of Democracy”, Chap. III, 23, *Modern Democracy*, Vol. I, Calcutta: The World Press Pvt. Ltd., 1962.

¹⁸Declaration of the Rights of Man made by the National Assembly of France, Cited in Bryce, Chapter V, *The Theoretical Foundations of Democracy*, 48-56, August 1791.

¹⁹The Constitution of India, art. 19(1)(a)

²⁰ John Stuart Mill, “On Liberty and Utilitarianism” 4 (Bantam Classic, New York, 2008)

right to express oneself freely in any medium and across all boundaries, free from outside interference such as censorship and fear of retaliation like threats and persecution; however, it is a complex right due to the fact that the right to free speech is not absolute and is subject to restrictions imposed by law.²¹ This implies that as long as one does not violate or infringe any laws or the rights of others, one is free to write or say whatever they like. Hate speech lies in a complex nexus with freedom of expression, individual, group and minority rights as well as concepts of dignity, liberty and equality.²²

Under international legal framework, speech is primarily governed by three key treaties, i.e. the International Covenant on Civil and Political Rights (ICCPR), the International Convention on the Elimination of All Forms of Racial Discrimination (ICERD) of 1965, and the Convention on the Prevention and Punishment of the Crime of Genocide of 1948. These outline the types of restrictions that states are legally obligated or encouraged to impose on speech. A couple of other international agreements including the European Convention on Human Rights, and the UDHR also guarantee the right to express themselves freely. These declarations specifically deal with the protection of freedom of speech and expression. The law of International human rights is less ambiguous regarding what constitutes restricted hate speech. These can be categorized into hate speech that states must prohibit and hate speech that states may prohibit”.²³ The severe effects of the very worst kinds of hate speech have been taken into consideration by the drafters of this section, and they also bind states to make these expressions illegal. An example of such a ban under international law is the direct and public incitement to genocide, which is prohibited by the 1948 Convention on the Prevention and Punishment of the Crime of Genocide, alongside with the prohibition of any crime of genocide itself.²⁴

Freedom of speech allows individuals to express their feelings to others, but this is not the only reason for protecting the freedom of speech. There are more reasons to protect these essential

²¹ Constitution, art. 19(2)- “Nothing in sub-clause (a) of clause (1) shall affect the operation of any existing law, or prevent the State from making any law, in so far as such law imposes reasonable restrictions on the exercise of the right conferred by the said sub-clause in the interests of the sovereignty and integrity of India, the security of the State, friendly relations with Foreign States, public order, decency or morality or in relation to contempt of court, defamation or incitement to an offence.”

²²Iginio Gagliardone, Danit Gal, Thiago Alves, Gabriela Martinez “Countering Online Hate Speech”, UNESCO Series On Internet Freedom, 2015, available at <http://unesdoc.unesco.org/ark:/48223/pf0000233231>, accessed on 20 June, 2024.

²³United Nations General Assembly, *Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression*, 18 UN Doc A/66/290 (10 August 2011). The Special Rapporteur has suggested three categories of hate speech to help identify appropriate and effective responses: (1) Hate speech that must be prohibited; (2) Hate speech that may be prohibited; (3) Lawful hate speech.

²⁴The Convention on the Prevention and Punishment of the Crime of Genocide, 9 December 1948.

liberties.²⁵ Freedom of speech is the foundation of any democratic society, serving as an essential right that empowers individuals to advocate for the realization of other rights, including fair trials, free elections, adequate living conditions, and more. “Liberty and equality are contemporary and not antithetical to each other. The intention of having the freedom of speech is not to disregard the weaker sections of society but to give them an equal voice. The intent of equality is not to restrain this liberty but to balance it with the necessities of a multicultural and plural world, provided such constraint does not unduly infringe on the freedom of expression. Thus, incitement to not only violence but also to discrimination has been recognized as a ground for interfering with freedom of expression.”²⁶ As discussed, currently, there is no universal category of hate speech that is uniformly penalized or prohibited. However, speech that poses threats, incites lawlessness, or motivates criminal activity can be prosecuted as a hate speech.

While examining the hate speech laws of India, the committee further recommended introducing new provisions within the penal code that specifically dealing with the provocation to violence in conjunction with the existing laws.²⁷ The harm caused by inciting to violence is worse than other forms of hate speech. The criminalization of speech is still a contentious topic, with legal experts disagreeing on what constitutes criminal speech, whether hate speech should only be prohibited in specific circumstances, and whether hate speech should always be dealt with under civil law rather than under criminal law. However, it is agreed that hate speech shown to be able to incite violence is a serious concern and needs to be addressed strongly to prevent any further damage. It is believed that using criminal sanction is the most effective and suitable way to curb hate.²⁸

BALANCING FREE SPEECH AND REGULATION

“Freedom of Speech doesn't justify online bullying. Words have power, be careful how you use them.”

²⁵Universal Declaration of Human Rights, art. 19, G.A. Res. 217A (III), U.N. GAOR, 3d Sess., U.N.Doc. A/810 (1948), which states, “Everyone has the right to freedom of opinion and expression; This right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers.”

²⁶ibid (n 11) 5.5.

²⁷ibid 11.

²⁸Katharine Gelber, “Speech Matters: Getting Free Speech Right” 84 (University of Queensland Press, Australia, 2011).

“There's no room for hate and violence in this world. We must learn to be more kind, compassionate, empathetic, and sympathetic to humanity.”

-Germany Kent

In case of *Rangarajan etc. v P. Jagjivan Ram*,²⁹ the SC held that to restrict free speech, a proximate and direct nexus must be established with any imminent danger to the community. Further, the Supreme Court in response to a public interest litigation seeking guidelines on hate speech during elections, observed that hate speech attempts to marginalize individuals based on their group membership.³⁰ This impacts people socially by diminishing their status, standing and acceptance. The Court also noted that existing laws in India are sufficient to tackle hate speech; the root of the problem lies not in the absence of laws, but in their efficacious execution.³¹

The regulation of hate speech in India, is primarily governed by the Indian Penal Code, 1860, now replaced by Bharatiya Nyaya Sanhita, 2023³² and various electoral and media laws. The BNS, 2023 criminalizes acts that incite communal disharmony or promote enmity or hatred among diverse groups. These laws aim to safeguard individual dignity and national unity by curbing divisive rhetoric.³³ The Bombay High Court, in the case of *Gopal Vinayak Godse*³⁴ provided a useful outline on the question of standards to assess the potential criminality of impugned content under section 153. Section 299 of BNS (295A of IPC) also penalizes acts done with deliberate and malicious intent to outrage the religious feelings of any class of citizens, as well as insults or attempts to insult their religion or beliefs. Similarly, other acts and speeches that incite violence are also made punishable.³⁵ The regulation of electoral conduct in India is governed by the Representation of the People Act, 1951³⁶ (hereinafter referred to as RPA, 1951), Articles 324–329 of the Constitution of India, and the Conduct of Election Rules, 1961. Hate speech during the elections is governed and regulated by the Model Code of Conduct (MCC) and the RPA, 1951. The RPA classifies hate speech during elections into 2 categories: electoral offences and corrupt practices. Relevant provisions include Section

²⁹[1989] 2 SCC 574

³⁰*Pravasi Bhalai Sangathan v. Union of India & Ors.*, AIR 2014 SC 1591.

³¹ *ibid.*

³² The Bharatiya Nyaya Sanhita, 2023 (Act 45 of 2023), hereinafter referred to as BNS.

³³ *ibid.*, s. 196, 197 (IPC, s. 153A, 153B).

³⁴*Gopal Vinayak Godse v. Union of India*, AIR 1971 Bom 56., 1961 SCR (3) 440

³⁵ BNS (n 32), s. 302, 353 (IPC s. 298, 505).

³⁶ Act No. 43 of 1951 w.e.f. 17 July 1951.

8 of RPA, 1951 which disqualifies individuals convicted of misuse of freedom of speech and expression from contesting elections. Additionally, Sections 123(3A) and 125 prohibit actions that promote enmity based on religion, race, caste, community, or language in the context of elections. These provisions criminalize hate speech by election candidates and mandate penalties for individuals- any person promoting feelings of enmity between different classes shall be punishable, with imprisonment for a term which may extend to three years, or fine, or with both.³⁷ Section 7 of The Protection of Civil Rights Act, 1955³⁸ penalizes incitement and encouragement of untouchability through words, either spoken or written, or by signs or by visible representations or otherwise. Section 3(g) of The Religious Institutions (Prevention of Misuse) Act, 1988³⁹ prohibits religious institution or their managers from permitting the use of premises under their control to promote or attempt to promote disharmony, enmity, hatred, or ill-will among different religious, racial, linguistic, regional groups, castes, or communities.

Similarly, the Information Technology Act, 2000 also addresses hate speeches and other abuse occurring online⁴⁰. The state recognized that increased use of the internet had led to new forms of crime, including offensive messages through communication services.⁴¹ Consequently, penal provisions were included in the IT Act, BNS, CrPC (now BNSS) and the Indian Evidence Act, 1872 (now BSA).⁴² Section 66A of the IT Act, introduced in 2008, was designed to address the dissemination of offensive and misleading messages. However, in the landmark case of *Shreya Singhal v. Union of India*⁴³ the Court differentiated between discussion, advocacy, and incitement, holding that the first two are the essence of Article 19(1). The Supreme Court struck down the controversial Section 66A in 2015, which penalized online hate speech offences. Following this, the Central Government constituted an Expert Committee⁴⁴ that proposed addressing online hate speech and spoofing separately through new sections under the IT Act. It recommended amendments to the IT Act to criminalize online content that ‘promotes ill will, hatred and enmity amongst communities, race, religions etc.’, similar to

³⁷The Representation of the People Act, 1951, (Act 43 of 1951), ss. 8, 123(3A), 125.

³⁸1st June, 1955, vide notification No. S.R.O. 1109E, dated the 23rd May, 1955, see Gazette of India, 1955, Extraordinary Part II, s. 3.

³⁹w.e.f. 26-05-1988

⁴⁰The Information Technology Act, 2000 (Act 21 of 2000) (hereinafter referred to as IT Act)

⁴¹The Information Technology (Amendment) Act 2008, Statement of Objects and Reasons

⁴²ibid.; The Information Technology Amendment Bill 2006

⁴³(2013) 12 SCC 73

⁴⁴Parliamentary Standing Committee on Home Affairs, 189th Report on ‘Action Taken by the Government on the recommendations/Observations Contained in the 176th Report on the Functioning of Delhi Police’, paras 3.5.5 (7 December 2015).

Sections 196 and 197 of the BNS.⁴⁵

While hate speech may not be explicitly addressed by these laws, the Constitution of India has been meticulously interpreted by the SC to confine these provisions under reasonable restrictions. As hate speech becomes increasingly prevalent, especially in era of social media, judicial interpretation of relevant legislation has become a crucial area of discussion in India. The dichotomy between safeguarding freedom of expression and regulating harmful speech is highlighted by various Supreme Court rulings, reflecting a cautious approach to the complexities of hate speech regulation. The Honorable Supreme Court has refrained from going beyond the purview of existing laws to penalize hate speech, as doing so would constitute ‘judicial overreach.’ The Court observed that effectively implementing existing laws would largely resolve the problem of hate speech to a great extent, as establishing a definite standard might lead to the curtailment of free speech.⁴⁶

In case of *Shaheen Abdulla v. UOI & Ors.*⁴⁷ the Supreme Court expressed alarm over the rising incidents of hate speech and instructed governments and police authorities to take suo motu action in such situations, without waiting for formal complaints. Furthermore, the Supreme Court ordered all States and Union Territories to take suo motu action to register FIRs against hate speech, regardless of the religion involved, without awaiting formal complaint.⁴⁸ In *Subramaniam Swamy v. Union of India*,⁴⁹ there had been pleas regarding the reasonableness of restrictions imposed by Sections 499-500 IPC on free speech, emphasizing that such restrictions should be narrowly tailored and should not be excessive, arbitrary or disproportionate.

The Honorable Supreme Court also made important observations on hate speech in case of *Amish Devgan vs. Union of India and others*,⁵⁰ which has generated discussion and controversy surrounding the concept of hate speech. The bench of Justices A.M. Khanwilkar and Sanjay Khanna discussed *the* need to distinguish between ‘free speech’ which includes the right to comment on, favor or criticize government policies and ‘hate speech’, which creates or

⁴⁵ibid 3.5.8.

⁴⁶ ibid 30.

⁴⁷ AIR 2022 SC 872

⁴⁸ “Supreme Court Directs All States/UTs To Register Suo Motu FIR Against Hate Speeches Irrespective Of Religion”, available at <https://www.livelaw.in/top-stories/supreme-court-directs-all-states-to-register-suo-motu-fir-against-hate-speeches-irrespective-of-religion-227444>, accessed on 18 Oct. 2023

⁴⁹[2016] 7 SCC 221

⁵⁰ 2020 SCC ONLINE SC 994

spreads hatred against targeted community or groups. The Court noted that the objective of criminalizing hate speech is to protect individual dignity and ensure political and social equality among different identities and groups, irrespective of caste, creed, religion, sex, gender identity, sexual orientation, linguistic preferences, etc. In this context, the Court clarified that hate speech does not refer to any particular level of honor or esteem as an individual, unlike defamation, which is individualistic.”⁵¹ Further, in *Arup Bhuyan vs State of Assam*⁵² the Court held that a mere act cannot be punished unless an individual resorts to violence or incites another person to violence. Over time, courts have consistently endeavored to define hate speech and delineate the types of speech that fall within its scope.

THE ROLE OF SOCIAL MEDIA IN AMPLIFYING HATE SPEECH

“Social media is not a media. The key is to listen, engage, and build relationships.”

-David Alston

We have transitioned from an era where individuals communicated primarily through oral means and limited to small groups, to one where they can connect instantaneously, with people across the globe through various digital platforms.⁵³ Today, nearly all generations are accustomed to the internet, and there has been a surge in the mobile usage with the emergence of smart phones. The use of social media in India in the last decade has witnessed a significant rise, further accelerated by affordable smart phones, cost-effective data plans, increased broadband penetration, and the expansion of internet access in regional and vernacular languages.

“The Internet has become a vital vehicle for promoting racism and intolerance. Hate speech on social media is rapidly increasing, reaching audience far beyond what extremist print media could ever achieve.”⁵⁴ Today, social media has the capacity to disseminate information and connect with a vast audience. While social networking websites were initially designed for

⁵¹ Sakshi Patil, “Hate Speech Repudiates Right To Equality In A Polity Committed To Pluralism: Supreme Court” (2020), available at <https://lawtimesjournal.in/hate-speech-repudiates-right-to-equality-in-a-polity-committed-to-pluralism-supreme-court/>, accessed on 21 May 2024.

⁵² [2011] 3 SCC 377

⁵³ Alice Tontodimamma, Eugenia Nissi, Annalina Sarra” Thirty years of research into hate speech: topics of interest and their evolution” *Scientometrics* 126, (2021).

⁵⁴ Council of Europe: European Commission Against Racism and Intolerance (ECRI), *Annual Report on ECRI's Activities 2014*, CRI (26, July 2015), available at <https://www.refworld.org/reference/annualreport/coecri/2015/en/105833>, accessed on 21 May 2024.

socializing, research, entertainment, building relationships, they are now misused to the fullest negative extent.⁵⁵ Online hate speech can lead to severe consequences, intimidating vulnerable communities into silence, particularly when it advocates incitement to hostility, discrimination or violence. If left unchecked, it could create an environment which undermines public discourse and can even harm those who do not use such platform. It is therefore important that States and companies address this problem with a determination to protect those at risk of being silenced and to promote open and rigorous debate on even the most sensitive issues in the public interest.”⁵⁶

According to a 2023 report by India Hate Lab, India recorded 668 documented hate speech events targeting Muslims. The report, titled '*Hate Speech Events in India*', noted a significant increase from 255 events in the first half of 2023 to 413 in the second half, marking a 62% rise.⁵⁷ Hate speech can persist on the internet for long periods, taking various formats across several platforms. The CEO of the Online Hate Prevention Institute commented that “the longer the content remains available, the more damage it can inflict on victims and empower the perpetrators.” Timely removal of such content can limit exposure. This is similar to picking up trash; it is not going to stop people from throwing it, but if you do not sort the issue, it will only become worse by piling up. It has further been argued that the internet grants individuals the ability to say offensive and vile things because they believe they will not be traced. This is what makes online hate speech so unique, as people feel much more comfortable expressing hate online compared to in real life, when they have to deal with the consequences of their words.⁵⁸

The spread of hate comments in social media is a challenging issue that needs to be addressed expeditiously. Anonymity poses a significant challenge in combating hate speech online. The emergence and spread of hate speech online is a growing phenomenon requiring concerted efforts to address its misuse and develop effective responses. The internet has facilitated an

⁵⁵Karuppannan Jaishankar, “Cyber Hate: Antisocial networking in the Internet”, Vol. 2 (2),16. July- December, 2008

⁵⁶United Nations General Assembly, *Report of the Special Rapporteur on the promotion and protection of the freedom of opinion and expression* UN Doc A/74/486 (9 October 2019)

⁵⁷ The Hindu Bureau, available at

<https://www.thehindu.com/news/national/75-of-hate-speech-events-in-bjp-ruled-states-report/article67888978.ece>, accessed on 15 March, 2024.

⁵⁸Gagliardone (n 22) 13-15.

explosion of extreme expression, exacerbated by the culture of anonymity that thrives online.⁵⁹

Social media is a powerful means of exercising one's right to freedom of speech and expression. However, it has also been used for illegal activities, prompting Governments to consider censorship. The challenge lies in balancing the need to eliminate objectionable and illegal content on social media while safeguarding civil rights of people. Proper regulation of social media is essential to protect and preserve the fundamental rights guaranteed under the Constitution of India. Instead of absolute freedom, which is not feasible; a balanced approach incorporating effective regulation, rather than outright censorship should be adopted. In 2021, the Ministry of Electronics and Information Technology notified Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021. These regulations primarily target social media, over-the-top (OTT) services, and digital news, aiming to provide users with grievance redressal mechanisms and ensure compliance with law, thereby regulating spread of hate through these channels. Specific legislation to regulate social media is necessary, although practical difficulties may arise in doing so. A thin line exists between enjoying one's right and the violating the rights of others. In social media, exercising freedom of speech and expression by one can lead to invasion of privacy and defamation of others. Therefore, it is crucial for the Government to protect both the right to freedom of speech and expression and the right to privacy in the realm of social media. To effectively address these challenges, the Government should collaborate with experts to develop comprehensive legislation that considers all facets of social media and possible misuse.

CONCLUSION AND SUGGESTIONS

From toddlers to elderly individuals, people of every age group are now exposed to smart phones and social media. These platforms have a wide and extensive reach throughout. Over time, the internet has increased the potential for harm with more options of mass communication leading to a surge in hateful messages and incitements to violence. Therefore, it is crucial to ensure that the harmful messages or hate messages are not disseminated frequently; they must be removed from social media as soon as possible. A viable solution needs to be presented by the authorities to minimize hate speech on social media platforms. Generally, recommendations are made to encourage self-regulation on the part of social media companies. However, no matter how much social media is trying to stop the spread of fake

⁵⁹Timothy Garton Ash, "Free Speech, Ten Principles for a connected World", 220 Atlantic Books, London, 2016.

news through its regulations, it is also our responsibility to address these issues earnestly. Simple and small actions, such as verifying the truthfulness of information before sharing it, can make a significant difference. We must also analyze the veracity of the articles and the videos before sharing it with someone else. Each of us needs to take responsibility for the content we create, consume and share. Primarily, we as citizens should learn to differentiate between free speech and hate speech. It is up to each one of us to refrain from spreading fake news or hate, while also being mindful not to hurt anyone's sentiments.

In light of the above discussion, it can be argued that rather than banning and censoring social media, appropriate regulation that protects user's rights is desirable. In India, hate speech, misinformation, and concerning rumors spread across several platforms have already resulted in violence and even murder. This serves as a dire warning, indicating that effective mechanisms must be implemented for its curtailment expeditiously. If we want our society to grow with love and compassion, it's absolutely necessary to foster understanding and acceptance. While dealing with the youth, it is essential to approach both the targets and the perpetrators of hate speech with empathy, as speaker may be unaware of the hateful implications of their language. While Indian penal laws seem to tackle hate speech with the required severity, confusion persists regarding the interpretation of what constitutes hate speech. This ambiguity arises because hate speech is not restricted to a particular definition or specific words; such restrictions could infringe upon the revered fundamental right to speech. Critics of such hate speech laws argue that they can infringe upon freedom of expression, and further interfering with other rights, such as equality, and freedom from discrimination and from violence. This ongoing debate highlights the challenge of balancing one's right to express thoughts with protecting community interests as well as preventing hate crimes. Efforts to navigate this complex issue are ongoing. Courts have tried on various occasions to demarcate a line between hate speech and other forms of speech that are often confused with it, but the dilemma continues to exist. It is evident that hate speech is an indefinable term, and existing laws are insufficient to address the issue effectively. Social media needs appropriate regulations to curb hate speech. It is possible that hate speech is still not fully understood by the relevant authorities. A clear definition and appropriate laws regarding hate speech could enable the judiciary to address it more effectively. A definition from layman's perspective can be proposed in the following words-

"Any communication, whether written or oral, that identifies patterns such as race, religion,

caste, region or other characteristics of like nature, and has the tendency to spread prejudice against any person or group, other than the speaker himself, can be described as Hate Speech.

Provided that it does not include communications intended to prevent any person or group from committing an offence and are made in good faith.

Explanation: Whether a communication has the tendency to spread prejudice is a matter of fact and shall be determined by reasonable group of individuals within society.”