

International Journal of Law Research, Education and Social Sciences

Open Access Journal – Copyright © 2025 – ISSN 3048-7501
Editor-in-Chief – Prof. (Dr.) Vageshwari Deswal; Publisher – Sakshi Batham



This is an Open Access article distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike 4.0 International (CC-BY-NC-SA 4.0) License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium provided the original work is properly cited.

From Expression to Harm: Addressing the Challenge of Online Hate Speech

Deeya Anil Dani^a Sanchya Prathish^b

^aSt. Joseph's College of Law, Bengaluru, India ^bRamaiah Institution of Legal Studies, Bengaluru, India

Received 03 March 2025; Accepted 03 April 2025; Published 07 April 2025

The rapid expansion of digital platforms has redefined the concept of freedom of expression, raising critical questions about its potential to foster online hate speech. This research investigates whether online hate speech is an inevitable consequence of unrestricted freedom of expression, particularly in global digital communication. The importance of democratic freedom is free expression; yet, unbridled exercise has helped vent vicious, discriminatory, and inflammatory speech that carries substantial ethical and societal burdens. Through critical examination, this study reviews international legal structures, ethical contracts, and technological moderation systems aimed at countering hate speech by interrogating individual rights against the interests of the greater collective goods. Further, it also looks into the socio-political and India's legal ramifications for internet hate speech and implications of unrestricted online discourse and the effectiveness of current measures in mitigating harm. The research underscores the need for a nuanced approach that balances the preservation of free speech with the imperative to protect individuals and communities from harm, where it doesn't slander or libel through the freedom of expression. It also looks at how other adjacent countries combat hate speech online and the extent to which they have implemented laws against it. The study concludes with a multi-faceted framework that brings together regulation, education, and technology to reduce online hate speech without undermining the principles of freedom of expression.

Keywords: *hate speech, freedom of speech, ipc, social media.*

INTRODUCTION: THE HATE SPEECH

“Darkness cannot drive out darkness; only light can do that. Hate cannot drive out hate; only love can do that.”

– Martin Luther King Jr.

These words are apt for the current world of digital landscape, which not only provides people such opportunities of expression but is also becoming increasingly frightening with a free space to let loose the worst forms of hate speech. The internet is our way of life now and has become an inevitable part of our lifestyle. Universal Declaration of Human Rights (1948)—Article 19 states that, ‘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’¹. But it is darkest in its potential use to hate speech, abuse of which is defined by the Cambridge Dictionary 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’². Hate speech is described differently by the United Nations, which seeks to incite hostility, discrimination, or even violence against anybody based on the characteristics of their identities.

The intersection of these two principles of freedom of expression and the curtailment of hate speech raises complex ethical and legal questions. This research seeks to explore whether hate speech is an inevitable byproduct of unrestricted free speech, offering insights into the societal, legal, and technological challenges in addressing this growing issue. As Voltaire once said, ‘I disapprove of what you say, but I will defend to the death your right to say it,’ a reminder of the delicate balance between liberty and responsibility in our modern discourse.

FREEDOM OF SPEECH AND EXPRESSION

“Without freedom of thought, there can be no such thing as wisdom and no such thing as public liberty without freedom of speech.”

¹ Universal Declaration of Human Rights 1948

² ‘Hate Speech’ (Cambridge Dictionary) <<https://dictionary.cambridge.org/us/dictionary/english/hate-speech>> accessed 11 January 2025

- Benjamin Franklin

India's history of freedom of speech and expression is complicated and predominantly shaped by colonial experiences as well as legislative modifications influenced by the significant events such as the French Revolution of 1789 and England's Bill of Rights 1689. This right is considered to be a form of liberty and personal autonomy that is practised in almost all democratic countries and others. It is known to be the 'Mother' of all other freedoms, as it is regarded as one of the crucial civil freedoms that are protected from state repression or limitation. In India, the right to freedom of speech and expression is protected under Article 19(1)(a), which is given solely to Indian citizens. Subject to the other provisions of this article, all citizens shall have the right – (a) to freedom of speech and expression.³

This provision grants every citizen the right to express words or writings to express their view and opinions. They can publish them without a fear of damage, but he or she will be held accountable for any harm caused or exercise of this right in the circumstances that the parliament shall determine. The essence of this right is for all Indian citizens to create their own opinions and express themselves as citizens, thereby participating actively in governmental activities and national events. It is important to know that this right is not absolute and that article 19(2)⁴ deals with the reasonable restrictions that the state imposes on the individuals such as the nation security, friendly relations with other countries, public order, contempt of court, morality, defamation, incitements and the sovereignty and integrity of India the eight grounds that one can be held accountable.

AMBIGUITY IN LEGAL FRAMEWORKS WITH CASE LAWS

The section deals with restriction that spread of hate speech offences as covered in the Indian penal code (IPC) the provisions against the hate speech are eloquent in three distant are Chapter I addresses 'Of Offences Relating to Religion,' Chapter II focuses on 'Of Offences Against the Public Tranquillity,; and Chapter III pertains to 'Of Criminal Intimidation, Insult and Annoyance.' The IPC provides the fundamental legal framework for the enforcement of substantive criminal law and is interpreted in conjunction with the Code of Criminal Procedure

³ Constitution of India 1950, art 19(1)(a)

⁴ Constitution of India 1950, art 19(2)

of 1973 (CrPC). The initial segment examines Section 153A of the IPC⁵, which makes it illegal to incite hostility between different groups based on religion and ethnicity. The second segment looks into Section 153B of the IPC, which prohibits statements and claims that could harm national unity. ⁶The third segment addresses Section 295A of the IPC, which penalises the act of vandalising places of worship or sacred items. The fourth segment focuses on Section 295A of the IPC, which makes it an offense to intentionally and maliciously act in a way that insults the religion or beliefs of any group, thereby provoking their religious sentiments. The fifth segment reviews Section 298 of the IPC, which outlaws expressions that could offend an individual's religious feelings. Finally, ⁷the sixth segment discusses Section 505 of the IPC, which criminalises the dissemination of particular statements, rumours or reports. Code of Criminal Procedure (CrPC) also provides for the arrest of individuals who are engaged in a cognizable crime that includes hate speech.⁸ For the protection of online forms of expression and views, the Indian Information Technology (IT) Act guides intermediaries like social media platforms to delete content that breaches the law within 36 hours of receiving a notification. Here are several regulations that limit individuals from promoting hate in any way that harms others and hold them accountable for their actions.

To understand the clear version of how these regulations work in India, let's look into some landmark cases where their judgments reflect the circumstances of hate speech in India:

Shreya Singhal v Union of India: ⁹ This important Supreme Court case struck down Section 66A of the Information Technology Act¹⁰. This section made online speech that was deemed offensive or annoying a crime. The court ruled the law was unclear, too large and violated the constitution's guarantee of free speech. The judgment strongly emphasised that the important right to freedom of expression includes the right to express dissent along with unpopular opinions. This ruling expanded the scope of free speech online. However, it also highlighted the challenge of dealing with harmful content without limiting individual liberties.

⁵ Indian Penal Code 1860, s 153A

⁶ Indian Penal Code 1860, s 295A

⁷ Indian Penal Code 1860, s 1505

⁸ Information Technology Act 2000, s 79

⁹ *Shreya Singhal v Union of India* (2015) SC 1523

¹⁰ Information Technology Act 2000, s 66A

P. Sukumar v State of Tamil Nadu:¹¹ The Madras High Court decisively determined that hate speech disseminated on social media platforms is unequivocally excluded from the protection afforded by Article 19(1)(a). The court made it clear that speech inciting violence, hatred, or discrimination disturbs public order. Therefore, the court can regulate this speech under Article 19(2), which permits reasonable restrictions on freedom of expression. The dedication of the judiciary to combating hate speech, while simultaneously upholding a meaningful number of constitutional values, was emphasised by this particular case.

Pravasi Bhalai Sangathan v Union of India & Ors (2014):¹² Upon examining the impact of hate speech on society, the Supreme Court recognised that it can spark violence and disrupt communal tranquillity. In this regard, the Court noted that existing laws in the Indian Penal Code, such as Sections 153A and 505, were insufficient to deal with the new forms which hate speech takes today, especially in cyberspace. The Court insisted that legislative intervention was required in order to strike an effective balance between regulation against freedom of speech. Importantly, however, without judicial decreeing law beyond its limits, Parliament needed to step in and fill this gap.

Bilal Ahmed Kaloo v State of Andhra Pradesh (1997):¹³ While hurting religious sentiments may be a ground for an offense under Sections 153B or 505 of the IPC, the Supreme Court clarified that it is not so if such a depiction or proof is used to incite hatred or violence. The court expressed that context and manner are important when adjudicating between mere offensiveness and imperilling public order. Failure to give importance to the ‘Blu’ test would be tantamount to curtailing free speech without reason.

Demonstrate the intricate balance between protecting freedom of speech and governing hate speech in India. These instances emphasise that although free expression is fundamental to democracy, it cannot be an unrestricted right when it provokes hatred, violence, or disturbs public order. The judiciary has repeatedly pointed out the inadequacy of current legal provisions, including Sections 153A, 295A, and 505 of the IPC, in tackling the evolving nature of hate speech, especially in the context of the digital age.

¹¹ *P. Sukumar v State of Tamil Nadu* (2019) WP No 16557

¹² *Pravasi Bhalai Sangathan v Union of India & Ors* (2014) SCC 477

¹³ *Bilal Ahmed Kaloo v State of Andhra Pradesh* (1997) SCC 431

CHALLENGES FACED TO CURB HATE SPEECH IN INDIA

Considering the existing regulations in India concerning the limitations of article 19(A) alongside with the new dimensions in the realm of freedom of speech and rise of online hate speech. India despite these laws still struggles to effectively tackle the issue of hate speech and mainly the online hate speech. those challenges are outlined as follows.

1. Gaps in defining hate speech in India: The definition of hate speech is not universal and lacks a clear reference to Indian law, making it difficult to interpret without a codified definition that relies on individuals to interpret and resulting in inconsistent enforcement. In the 267th Report of the Law Commission of India, hate speech is stated as an incitement to hatred primarily against a group of persons defined in terms of race, ethnicity, gender, sexual orientation, religious belief and the like.¹⁴ With regards to this statement, it lacks the contextual considerations and the clear thresholds for incitements in ambiguity of defining the level of harm or damage to be considered to classify it as hate speech. Additionally, considering a country like India with the largest population country the existing diversity in race, language, culture, customs and religion becomes difficult to recognise due to this diversity, individuals' sentiments and perspectives often differ among the citizens, which makes it hard to acknowledge the hate speech.

2. Regulatory Gaps in Addressing Hate Speech in India: Hate speech in India confronts several challenges in a society that is committed to diversity and harmony. Striking a balance between the constitutional guarantee of freedom of speech, Article 19(1)(a), and the necessity of curbing divisive or hate speech poses an ongoing challenge. Existing legal provisions (Sections 153A, 295A and 505 of the Indian Penal Code) remain vague and provide scope for irregular application as well as abuse. The internet or internet-enabled platforms not only amplify hate speech but also present novel regulatory conundrums: algorithms that facilitate the promotion or propagation of hateful content and anonymous postings that foster impunity. Politicisation ensures that such utterances, primarily within the realm of political discourse, become invisible rhetoric meant to sway voters, thereby increasing communal polarisation. Law enforcement agencies lack both resources and expertise to better regulate this space; a lack of sufficient

¹⁴ 'Hate Speech' (*Drishti IAS*, 07 October 2023) <<https://www.drishtias.com/daily-updates/daily-news-analysis/hate-speech-5>> accessed on 12 January 2025

policing capacity leads to ineffective regulation online. Differential application also lends itself to loss of credibility for both law enforcement agencies (LEAs) as well as the judiciary about their treatment of influential individuals or groups who may use such language. Addressing these issues requires clear legislative guidelines on what constitutes hate speech, active pursuit by LEAs in cases where a prima facie case exists, along with social determination at all levels against using language aiming to incite hatred while reaffirming the principles enshrined in our Constitution.

3. Conflict between Freedom of Speech and Regulations enforced by the state:

Article 19(1)(a) guarantees freedom of expression and speech but as this right is not absolute the Article 19(2) provides us the reasonable restrictions there are eight reasonable grounds as mentioned before. However, the prohibition of expression on these grounds has produced undesirable consequences for India, given its political, cultural, and religious diversity and sensitivities and the ever-present possibility of slighting some cherished beliefs or views. Despite the governmental claims of democratic validity, a culture of intolerance towards criticism has been incubating among the paranoid nationalist and religious factions in the country, as they fail to accurately distinguish between those that threaten public order, decency, or morality and those that threaten their political or personal sensitivities. There have been several instances where individuals, groups, and media organisations have faced intimidation, harassment, and violence for expressing dissenting views.¹⁵ And to curb any intermediate online violence in India has started internet shutdowns seemingly orchestrated by the government to silence opposition and control public sentiment, and detailed scrutiny.¹⁶ For example, in early 2023, Manipur witnessed appalling incidents where women were subjected to unspeakably heinous acts, including women being paraded naked¹⁷, gang raped and brutally murdered¹⁸ and this was occurred during the total blackout of internet facilities in the regions as a result there was

¹⁵ Subhajit Basu and Shameek Sen, 'Silenced Voices: Unravelling India's Dissent Crisis Through Historical and Contemporary Analysis of Free Speech and Suppression' (2013) 33(1) Information and Communications Technology Law <<https://doi.org/10.1080/13600834.2023.2249780>> accessed on 12th January 2025

¹⁶ Bajoria, 'No Internet Means No Work, No Pay, No Food' (*Human Rights Watch*, 14 June 2023) <<https://www.hrw.org/report/2023/06/14/no-internet-means-no-work-no-pay-no-food/internet-shutdowns-deny-access-basic>> accessed 12 January 2025

¹⁷ 'Human Rights Commission Issues Notice to Manipur on Women Being Paraded Naked' *India Today* (21 July 2023) <<https://www.indiatoday.in/india/story/human-rights-commission-issues-notice-to-manipur-on-women-being-paraded-naked-2409540-2023-07-21>> accessed 12 January 2025

¹⁸ Saraswat Kashyap, 'Manipur Violence: 2 Women Gang-Raped, Killed on Same Day of Video Incident' *India Today* (24 July 2023) <<https://www.indiatoday.in/india/story/manipur-violence-alleged-rape-abduction-murder-of-two-women-2410387-2023-07-22>> accessed 12 January 2025

suppression of this news for nearly 3 months¹⁹. The misuse of limitations to stifle opposition and control narratives presents a serious risk to democratic principles. Incidents of Manipur violence during these complete blackouts demonstrate how these measures can suppress essential information and postpone justice.

4. Influence of Social Media on Public Opinion: Social media platforms are making huge gains in shaping opinions, influencing decisions, and connecting people globally. With over half the Indian population active on social media, their potential to influence societal trends, public opinion, and even political discourse turns out to be monumental. However, the absence of stringent regulations to govern online behaviour has turned them into hotbeds for hate speech, misinformation, and divisive discourse. One of the defining moments of the Indian social media arena came with the banning of TikTok in June 2020. Though the basis for banning this TikTok application was largely because of allegations against data leaks by Chinese-owned apps, this application got continuously scrutinised for permitting a larger chunk of uncontrolled, malicious content, including hate speech, communal propaganda, and abusive videos. In 2011 Indian government required social media platforms such as Google, Facebook, and Yahoo to pre-screen user-generated content, especially to address hate speech and harmful materials.²⁰ This initiative was spearheaded by Communications Minister Kapil Sibal as it was triggered online and perceived as an infringement on free expression. ²¹Despite these regulatory measures in India 2019 elections hate speech aimed at female politicians surged with 13.8% of tweets directed at 95 women politicians being abusive and often sexual, sexist, and caste related insults from the social media platforms and the prohibition of apps like PUBG Mobile and UC Browser in 2020 due to their roles in disseminating hate speech and harmful content highlighted the shortcomings of these responsive actions. Despite these initiatives, it is evident that the existing measures are inadequate to tackle the comprehensive extent of online hate speech. Social media platforms remain one of the media for disseminating misinformation and hatred, especially during times of social and political sensitivity.

¹⁹ Srinivas Kodali, 'We Can't Look Away from Internet Shutdowns in Manipur' *The Wire* (20 July 2023) <<https://thewire.in/rights/we-cant-look-away-from-internet-shutdowns-in-manipur>> accessed 12 January 2025

²⁰ Heather Timmons, 'India Asks Google, Facebook to Screen User Content' *The New York Times* (05 December 2011) <<https://archive.nytimes.com/india.blogs.nytimes.com/2011/12/05/india-asks-google-facebook-others-to-screen-user-content/>> accessed 13 January 2025

²¹ 'Elections once again - and it's press play on sexist slurs against women politicians' *The Economic Times* (07 April 2024) <<https://economictimes.indiatimes.com/news/elections/lok-sabha/india/elections-once-again-and-its-press-play-on-sexist-slurs-against-women-politicians/articleshow/109108303.cms>> accessed 13 January 2025

5. Global Perspectives: In 2019, the United Nations (UN) adopted a Strategy and Plan of Action on Hate Speech, recognising the potential that hate speech has to provoke violence and destroy societies²². Speech that attacks or uses language that is discriminatory to certain individuals or groups on the grounds or basis of their identity, such as race, religion, gender, or ethnicity, is defined broadly by the UN as ‘communication which attacks or uses prejudiced language aimed at persons or groups based on their identity.’²³ The initiative addresses the escalating supply of hate speech in triggering conflicts such as the Rwandan Genocide and the subsequent role of digital platforms in propagating narratives based on hate.

It argues for a balance between the right to combat hate speech and the right to freedom of expression. It recruits governments, tech companies, and individuals to work together in combating hate speech, which means legally enabling clear frameworks, digital literacy programs, and counter-narratives in favour of inclusion and respect.

The UN also points out that social media algorithms are increasingly being used for this, according to reports. It sometimes propagates inflammatory content just for the sake of garnering engagement, thereby further spreading toxic rhetoric. This strategy aims at fostering accountability, promoting tolerance, and strengthening global actions in developing peaceful and inclusive societies, while at the same time not trampling the right to freedom of expression.

PROPOSED SOLUTION AND SUGGESTIONS

The RFF is an empowering strategy to help deal with hate speech by emphasising maintaining a proper balance between freedoms of expression and social peace. This framework envisages a heavy onus of responsibility on social media platforms, along with the retraction of certain legal and ethical limits applied to online discourse.

1. The Responsibility of Algorithm: Social media platforms should completely reform their algorithms to support messages of inclusion and critical engagement, not incendiary posts. Severely, the state shall impose heavy fines on those platforms that profit from hate messages.

²² ‘United Nations Strategy and Plan of Actions on Hate Speech’ (*United Nations*, 18 June 2019) <<https://www.un.org/en/genocideprevention/documents/UN%20Strategy%20and%20Plan%20of%20Action%20on%20Hate%20Speech%2018%20June%20SYNOPSIS.pdf>> accessed 13 January 2025

²³ *Ibid*

2. An Online Speech License: The RFF introduces somewhat controversially a digital speech license wherein users would have to write a mandatory test on digital ethics and hate speech awareness before being allowed to use online media. Repeat offenders will be subject to graded penalties, such as partial and complete bans.

3. Real-Time Monitoring & AI: This independent international organisation, which would be funded by contributions from governments and technology firms, will oversee real-time monitoring on the basis of AI tools whose expertise is programmed to detect hate speech quickly so as to facilitate immediate and swift removal of hate posts that disrupt the peace.

4. Community Justice: Communities that are victims of hate speech should impose punishment on the offenders, thus bringing empowerment and accountability.

5. Education First: It is legally required that every school program under construction must incorporate modules on digital ethics and anti-hate speech awareness.

Some adversaries might view RFF as an infringement on freedom of expression. However, this framework seeks to enhance responsible exercises of expression rather than cause anarchy, for imposition of untrammelled freedom without responsibility will ruin a democracy.

CONCLUSION

In summary, the issue of hate speech in the digital era is a complex challenge at the intersection of freedom of speech, societal cohesion, and legal regulations. While the right to speech and expression remains a right, its abuse through hate speech, particularly in the digital realm, poses threats to social peace, personal dignity, and democratic values. India's legal framework is strong in certain respects; it remains inadequate to fully address the nuances and scale of online hate speech. There is an immediate necessity for more stringent regulatory measures and improved enforcement mechanisms to hold perpetrators accountable while safeguarding individual freedoms. Adopting a more active and inclusive strategy, like establishing dedicated portals and the RFF solutions, reporting hate speech practices and laws, accommodates the complexities of social media. The balance between freedom of speech and the protection of individuals from harm is delicate, but achievable. As society continues to evolve, so too must our legal systems, ensuring that the digital space remains a platform for healthy expression rather than a breeding ground for hate.