BHARGAVA Insights Multidisciplinary Research Journal
ISSN:
Volume: 01, Issue: 01, April- June 2026
https://bhargavafoundation.in/

HATE SPEECH AND ONLINE HARASSMENT:

A LEGAL AND SOCIAL CHALLENGE IN THE DIGITAL AGE

BIMRJ
ISSN
Double-Blind Peer Reviewed
Open Access Quarterly Journal

Ms Anupam Sinha,

Assistant Professor, Amity Law School,

Amity University, Patna

ABSTRACT

Online toxic discourses could result in conflicts between groups or harm to online communities. Hate speech is complex and multifaceted harmful or offensive content targeting individuals or groups. Existing literature reviews have generally focused on a particular category of hate speech, and to the best of our knowledge, no review has been dedicated to hate speech datasets. This paper systematically reviews textual hate speech detection systems and highlights their primary datasets, textual features, and machine learning models. The results of this literature review are integrated with content analysis, resulting in several themes for 30 relevant papers. This study shows several approaches that do not provide consistent results in various hate speech categories.

***Corresponding Author**
Ms Anupam Sinha, Assistant Professor, Amity Law School,

Amity University, Patna

✉ **anupam_sinha123@yahoo.co.in**

Article Info
Received: 28th January 2026
Final Accepted: 30th March 2026
Published: 9th April 2026

The most dominant sets of methods combine more than one deep learning model. Moreover, the analysis of several hate speech datasets shows that many datasets are small in size and are not reliable for various tasks of hate speech detection. Therefore, this study provides the research community with insights and empirical evidence on the intrinsic properties of hate speech and helps communities identify topics for future work.

Keywords– Hate speech, online harassment, cyberbullying, Anonymity, Phishing, Impersonation.

INTRODUCTION

"Addressing hate speech does not mean limiting or prohibiting freedom of speech. It means keeping hate speech from escalating into something more dangerous, particularly incitement to discrimination, hostility and violence, which is prohibited under international law."

— United Nations Secretary-General António Guterres, May 2019

In the digital age, communication has become faster, easier, and more widespread. However, the proliferation of social media platforms and online forums has also created new spaces for harmful behavior. Among the most alarming are hate speech and online harassment, both of which have emerged as significant threats to digital safety and civil discourse. While the internet provides a platform for free expression, it has also enabled individuals and groups to spread discrimination, incite violence, and target others with abuse, often anonymously and without immediate consequence. Today, when the internet is present in every aspect of our lives (education, information, shopping, etc.), young people are interested in the latest trends. Children use the internet and mobile phones for easy access to information, for better and broader communication, and for socialization; and all of this can happen anytime, anywhere. However, these are all advantages of technology, but along with this come some risks. Students attending colleges and universities in this strange country today, instead of using the internet to search for new ideas and their information is lost, try to keep up with pornography, cyberbullying, harassment, etc. Today.

In addition, young people can use the Internet and mobile phones to send embarrassing images or messages, harass someone, spread false rumors about someone, and engage in acts that are derogatory about their physical appearance with the intention of harming their physical appearance. These online actions are called cyberbullying, which is defined as "repeated insults or antisocial behaviour by a group or individual electronically against a defenceless victim."

Therefore, cyberbullying is a type of bullying that occurs using electronic technology. The term cyberbullying can be used by many different names: cyberbullying, e-bullying, cyberbullying, cyberbullying, text bullying, SMS bullying, mobile phone bullying, cyberbullying, and internet bullying. In this article, we use the term "cyberbullying" to refer to harassment of others using modern electronic technologies, especially the Internet and mobile phones.

INTERNATIONAL PROSPECTIVE

Hate speech and online harassment are not confined to national borders—they are global phenomena that transcend linguistic, cultural, and political boundaries. As internet access becomes more widespread, so does the potential for abuse and harm. The international community has increasingly recognized the severity of this issue and has taken steps to address it through legal frameworks, multilateral cooperation, and the promotion of digital rights. However, significant differences in legal standards, cultural values, and enforcement mechanisms present challenges to a unified global approach.

United Nations and International Human Rights Law

At the core of the international response is the United Nations, particularly through the International Covenant on Civil and Political Rights (ICCPR). Article 19 of the ICCPR protects freedom of expression, while Article 20 specifically prohibits advocacy of hatred that constitutes incitement to discrimination, hostility, or violence. The challenge lies in balancing these two provisions—ensuring that freedom of expression is not curtailed unnecessarily while also protecting individuals from harmful speech.

The Rabat Plan of Action, issued by the UN Office of the High Commissioner for Human Rights (OHCHR), offers guidelines for implementing restrictions on hate speech in a manner consistent with international human rights norms. It introduces a six-part threshold test for determining whether certain speech should be legally limited, considering factors such as the speaker's intent, the content, and the likelihood of harm.

Europe: A Regulatory Leader

European countries have been at the forefront of regulating hate speech and online abuse. The European Union's Framework Decision on Racism and Xenophobia (2008) requires member states to criminalize public incitement to violence or hatred based on race, religion, ethnicity, or national origin. In 2022, the EU passed the Digital Services Act (DSA), which imposes strict

obligations on large online platforms to remove illegal content, including hate speech and online harassment, in a timely manner.

Additionally, many European countries have their own national laws:

- Germany's Network Enforcement Act (NetzDG) requires social media platforms to take down hate speech within 24 hours of notification.
- France's Avia Law, though partially struck down, aimed to tackle online hate by holding platforms accountable.

These laws demonstrate Europe's proactive stance but have also raised concerns about overreach and potential censorship, particularly in politically sensitive contexts.

United States: Strong Free Speech Protections

In contrast, the United States takes a more liberal approach, prioritizing free speech under the First Amendment. As a result, hate speech is generally protected unless it meets the criteria for incitement to imminent violence or constitutes a "true threat." While this framework protects political dissent, it also allows significant amounts of harmful speech to flourish online.

Despite the constitutional limits, the U.S. has laws against cyberstalking, doxxing, and harassment, often implemented at the state level. Federal initiatives like the Online Harassment Task Force, formed under the Department of Justice, aim to address the rise in digital abuse, particularly against women and minorities.

Global South: Emerging Responses and Challenges

Countries in the Global South, such as India, Brazil, and South Africa, face distinct challenges in tackling hate speech and online harassment. These include:

- Weak enforcement of existing laws
- Limited digital literacy
- Politicization of hate speech regulation
- Language diversity and technical barriers in content moderation

India, for instance, has laws like Sections 153A and 295A of the Indian Penal Code to deal with hate speech, and the Information Technology Rules, 2021, which impose responsibilities on

intermediaries. However, enforcement is uneven, and laws are sometimes misused to suppress dissent.

Toward a Harmonized Global Approach

Despite regional differences, there is a growing recognition of the need for international cooperation. Organizations like the Council of Europe, UNESCO, and Internet Governance Forum (IGF) promote global dialogue on balancing freedom of expression with the need to combat online hate. There is also increasing pressure on tech giants—many of which operate globally—to adopt consistent, transparent content moderation practices across jurisdictions.

In conclusion, while no single model fits all, the international community must strive for harmonized standards, rights-based regulation, and collaborative enforcement to ensure that the internet remains a safe space for all.

INDIAN PROSPECTIVE

India, with its vast population and rapidly expanding internet user base, has witnessed a significant surge in online interactions over the past decade. While this digital growth has enabled communication and innovation, it has also exposed users to the darker aspects of the internet—particularly hate speech and online harassment. These forms of abuse have become increasingly common on social media platforms, messaging apps, and online forums, raising urgent questions about digital safety, freedom of expression, and the effectiveness of legal frameworks.

Legal Frameworks Addressing Hate Speech and Online Harassment

India's approach to regulating hate speech and online harassment is primarily rooted in the Indian Penal Code (IPC) and the Information Technology (IT) Act, 2000.

Indian Penal Code (IPC)

Several sections of the IPC criminalize hate speech:

- Section 153A: Punishes promoting enmity between different groups on grounds of religion, race, language, etc., and acts prejudicial to maintaining harmony.
- Section 295A: Penalizes deliberate and malicious acts intended to outrage religious feelings.

- Section 505(1) & 505(2): Address statements that cause public mischief or incite communities to commit offenses against each other.

These provisions are often invoked in cases of communal hate speech or inflammatory content shared online. However, the broad and vague wording of some of these laws has led to criticisms of misuse and selective enforcement.

Information Technology Act, 2000

The IT Act, India's primary cyber law, contains specific provisions to address online abuse:

- Section 66A (now struck down): Previously criminalized sending offensive messages electronically, but was declared unconstitutional by the Supreme Court in *Shreya Singhal v. Union of India (2015)* for violating freedom of speech.
- Section 66E: Penalizes the violation of privacy through capturing, publishing, or transmitting private images without consent.
- Section 67 and 67A: Punish publishing or transmitting obscene and sexually explicit content online.

The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 further outline the responsibilities of online platforms in moderating harmful content and responding to user complaints. These rules mandate due diligence by intermediaries, grievance redressal mechanisms, and traceability of originators of offensive content in certain cases.

Societal and Political Dimensions

India's complex social fabric, marked by diversity in religion, caste, language, and region, creates fertile ground for identity-based online hate. Political discourse has also become increasingly polarized, with hate speech being weaponized for electoral gains or to silence dissenting voices.

Women, journalists, activists, and minority communities are disproportionately targeted by troll armies and coordinated harassment campaigns. The anonymity of digital platforms often emboldens perpetrators, while victims struggle to receive timely support from law enforcement.

Challenges and the Way Forward

Key challenges in addressing hate speech and online harassment in India include:

- Ambiguous laws that can be misused to suppress legitimate expression
- Lack of digital literacy among users
- Inadequate enforcement capacity of cyber cells and police
- Limited accountability of social media platforms

To move forward, India must:

- Update and clarify existing legal definitions
- Enhance digital literacy and awareness campaigns
- Train law enforcement in cybercrime investigation
- Hold tech companies accountable through transparent and rights-based regulations

India's battle against online hate and harassment is far from over. A careful balance must be struck between protecting free speech and curbing harmful online behavior to ensure a safer digital space for all.

JUDICIAL PROSPECTIVE

The Indian judiciary has played a pivotal role in shaping the legal landscape surrounding hate speech and online harassment. Through landmark judgments and evolving interpretations, courts in India have sought to balance freedom of speech (Article 19(1)(a)) with reasonable restrictions (Article 19(2)) on the grounds of public order, decency, morality, and the sovereignty and integrity of India.

Shreya Singhal v. Union of India (2015)

One of the most significant judgments in this context is the Shreya Singhal case, where the Supreme Court struck down Section 66A of the Information Technology Act, 2000 for being vague and overbroad. Section 66A criminalized sending "offensive" messages through communication services, which the Court held as a violation of free speech under Article 19(1)(a). The judgment underscored that mere annoyance or inconvenience cannot be grounds for curbing speech and that restrictions must align with Article 19(2).

This case became a cornerstone in delineating the limits of state intervention in online expression, setting a strong precedent for protecting digital rights.

Pravasi Bhalai Sangathan v. Union of India (2014)

In this case, the Supreme Court addressed the issue of hate speech and urged Parliament to enact stricter laws. While the Court refrained from issuing specific guidelines, it recognized the growing menace of hate speech, especially in the digital domain, and called upon lawmakers to frame a clear legal framework to regulate it without infringing on constitutional rights.

Amish Devgan v. Union of India (2020)

The Court in this case upheld multiple FIRs against a TV anchor for allegedly hurting religious sentiments during a broadcast. The judgment clarified that hate speech must be assessed in context, including the speaker's intent and the potential impact on public order. It reinforced the principle that freedom of speech is not absolute, and speech that incites violence or disrupts harmony can be lawfully curtailed.

Judicial Trends and Observations

- Indian courts have generally favored a case-by-case approach when it comes to online speech.
- They have emphasized proportionality and contextual interpretation while assessing whether a particular expression qualifies as hate speech.
- Courts have also acknowledged the need to update existing laws to tackle emerging challenges posed by social media, fake news, and trolling.

The judiciary has been a guardian of constitutional freedoms while also recognizing the growing threat of digital abuse. However, the absence of specific legislation on hate speech has left much of the burden on judicial discretion, highlighting the need for comprehensive statutory reform that aligns with democratic values and evolving technologies.

CONCLUSION

Hate speech and online harassment represent a profound challenge in the digital age, threatening the core values of dignity, equality, and freedom of expression. As the internet continues to evolve into a primary platform for public discourse, the potential for misuse

through targeted abuse, communal incitement, and gendered or identity-based harassment has become increasingly evident. These forms of online abuse not only cause psychological harm but also erode democratic participation and social harmony.

Globally, countries have adopted varied legal responses—ranging from strict regulatory regimes in Europe to free speech–oriented approaches in the United States. India, with its unique socio-political landscape and diverse population, faces a complex interplay of constitutional rights, societal fault lines, and technological advancements. While Indian laws such as the Indian Penal Code and the Information Technology Act provide some mechanisms to address online abuse, challenges remain in the form of vague legal provisions, inconsistent enforcement, and limited digital literacy.

The Indian judiciary has acted as a crucial mediator in balancing fundamental rights with the need for regulation. Landmark judgments like *Shreya Singhal v. Union of India* have emphasized the importance of clarity in legal provisions and the protection of free speech, while others have recognized the urgency of combating hate-driven and harmful online expression.

However, judicial activism alone cannot address the multifaceted nature of hate speech and online harassment. What is needed is a holistic approach—one that combines clear and updated legislation, robust content moderation by tech platforms, public awareness campaigns, and stronger institutional capacity for cybercrime investigation and redressal. International cooperation and shared regulatory norms are also essential, given the borderless nature of digital communication.

Ultimately, the goal must be to create a digital environment where freedom of expression thrives, but not at the cost of human dignity or safety. Protecting users—especially women, minorities, and marginalized voices—requires coordinated efforts across legal, technological, and social spheres. As societies grow increasingly interconnected, the fight against hate speech and online harassment is not just a legal necessity, but a moral and democratic imperative.

SUGGESTIONS

To effectively combat hate speech and online harassment while safeguarding freedom of expression, a multi-stakeholder and balanced approach is essential. The following suggestions are aimed at strengthening the legal, technological, and social response to this growing problem:

1. Enact Comprehensive Legislation

- India and other countries should consider enacting specific, clear, and narrowly defined laws to address hate speech and online harassment.
- These laws must comply with constitutional principles and international human rights norms, ensuring that they are not misused to stifle dissent or criticism.

2. Enhance Digital Literacy and Public Awareness

- Launch nationwide campaigns to educate users—especially youth—about responsible online behavior, cyberbullying, and the legal consequences of online abuse.
- Include digital safety education in school and college curricula to build resilience and awareness from an early age.

3. Strengthen Law Enforcement Capacity

- Equip cybercrime units with advanced tools and training to effectively investigate and prosecute online abuse.
- Establish fast-track mechanisms for victims to report hate speech and harassment, with assured protection and psychological support.

4. Hold Social Media Platforms Accountable

- Platforms like Facebook, X (Twitter), Instagram, and YouTube must implement stronger content moderation policies, especially in regional languages.
- Require platforms to publish transparency reports, detailing takedown requests, algorithmic moderation, and complaint resolution statistics.

5. Empower Victims and Support Civil Society

- Set up helplines and legal aid services specifically for victims of online abuse, especially women, LGBTQ+ individuals, and marginalized communities.
- Encourage collaboration with NGOs and civil rights groups to track online hate trends, support victims, and advocate for reform.

6. Encourage International Collaboration

- Develop cross-border frameworks for information sharing and enforcement, especially for hate speech that originates or is coordinated internationally.
- Support global norms for platform regulation, content moderation standards, and user data protection.

7. Foster Ethical Tech Development

- Promote the use of AI-driven moderation tools with built-in safeguards to detect hate speech without infringing on legitimate speech.
- Encourage ethics training for developers and stronger human oversight in algorithmic decision-making.

These suggestions, if implemented collectively, can contribute to a safer, more inclusive, and respectful digital environment, where freedom of speech coexists with accountability and human dignity.

REFERENCES

- Shreya Singhal v. Union of India, (2015) 5 SCC 1.
- Amish Devgan v. Union of India, (2021) 1 SCC 1.
- Pravasi Bhalai Sangathan v. Union of India, (2014) 11 SCC 477.
- Information Technology Act, 2000, No. 21, Acts of Parliament, 2000 (India).
- Indian Penal Code, 1860, No. 45, Acts of Parliament, 1860 (India).
- United Nations, *International Covenant on Civil and Political Rights*, 1966, available at: https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights (last visited May 20, 2025).
- Office of the United Nations High Commissioner for Human Rights, *Rabat Plan of Action on the Prohibition of Advocacy of National, Racial or Religious Hatred*, 2012, available at: https://www.ohchr.org/en/documents/publications/rabat-plan-action (last visited May 20, 2025).
- European Union, *Digital Services Act*, Regulation (EU) 2022/2065, OJ L 277, 27.10.2022, p. 1–102.

- David Kaye, *Speech Police: The Global Struggle to Govern the Internet* (Columbia Global Reports, 2019).
- Ranjan, Radha, Anand Shyam Kumar , Meena Bheem Singh (2023). virtual meetings under attack: assessing the legal and security risks of zoom bombing in the digital era. ShodhKosh: Journal of Visual and Performing Arts July-December 2023 4(2), 1256–1263. https://doi.org/10.29121/shodhkosh.v4.i2.2023.3047
- Ranjan, M. R. (2024). Digital Personal Data Protection Act 2023: Safeguarding Your Online Identity. *Sustainable Development Goals & Business Sustainability*, 108.
- Ranjan, R., Bhaumik, J., & Patel, A. (2024). Regulating artificial intelligence in legal practice: an accountability framework. *blockchain and ai in business*, 131.
- Indian Ministry of Electronics and Information Technology, *Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021*, Gazette Notification G.S.R. 139(E), dated 25.02.2021.
- Singh Pallavi, Ranjan, Radha (2023). Cyber Crime Against Women In Cyber Space: A Critical Analysis of Indian Legislations. Kanpur Philosopher UGC CARE Listed Journal, ISSN No- 2348-8301., 10(1(A), 79–85.