

INTERNATIONAL JOURNAL FOR LEGAL RESEARCH AND ANALYSIS



Open Access, Refereed Journal Multi-Disciplinary
Peer Reviewed

www.ijlra.com

DISCLAIMER

No part of this publication may be reproduced or copied in any form by any means without prior written permission of Managing Editor of IJLRA. The views expressed in this publication are purely personal opinions of the authors and do not reflect the views of the Editorial Team of IJLRA.

Though every effort has been made to ensure that the information in Volume II Issue 7 is accurate and appropriately cited/referenced, neither the Editorial Board nor IJLRA shall be held liable or responsible in any manner whatsoever for any consequences for any action taken by anyone on the basis of information in the Journal.

Copyright © International Journal for Legal Research & Analysis

EDITORIALTEAM

EDITORS

Dr. Samrat Datta

Dr. Samrat Datta Seedling School of Law and Governance, Jaipur National University, Jaipur. Dr. Samrat Datta is currently associated with Seedling School of Law and Governance, Jaipur National University, Jaipur. Dr. Datta has completed his graduation i.e., B.A.LL.B. from Law College Dehradun, Hemvati Nandan Bahuguna Garhwal University, Srinagar, Uttarakhand. He is an alumnus of KIIT University, Bhubaneswar where he pursued his post-graduation (LL.M.) in Criminal Law and subsequently completed his Ph.D. in Police Law and Information Technology from the Pacific Academy of Higher Education and Research University, Udaipur in 2020. His area of interest and research is Criminal and Police Law. Dr. Datta has a teaching experience of 7 years in various law schools across North India and has held administrative positions like Academic Coordinator, Centre Superintendent for Examinations, Deputy Controller of Examinations, Member of the Proctorial Board



Dr. Namita Jain

Head & Associate Professor

School of Law, JECRC University, Jaipur Ph.D. (Commercial Law) LL.M., UGC -NET Post Graduation Diploma in Taxation law and Practice, Bachelor of Commerce.

Teaching Experience: 12 years, AWARDS AND RECOGNITION of Dr. Namita Jain are - ICF Global Excellence Award 2020 in the category of educationalist by I Can Foundation, India. India Women Empowerment Award in the category of "Emerging Excellence in Academics by Prime Time & Utkrisht Bharat Foundation, New Delhi. (2020). Conferred in FL Book of Top 21 Record Holders in the category of education by Fashion Lifestyle Magazine, New Delhi. (2020). Certificate of Appreciation for organizing and managing the Professional Development Training Program on IPR in Collaboration with Trade Innovations Services, Jaipur on March 14th, 2019



Mrs.S.Kalpana

Assistant professor of Law

Mrs.S.Kalpana, presently Assistant professor of Law, VelTech Rangarajan Dr.Sagunthala R & D Institute of Science and Technology, Avadi. Formerly Assistant professor of Law, Vels University in the year 2019 to 2020, Worked as Guest Faculty, Chennai Dr.Ambedkar Law College, Pudupakkam. Published one book. Published 8Articles in various reputed Law Journals. Conducted 1Moot court competition and participated in nearly 80 National and International seminars and webinars conducted on various subjects of Law. Did ML in Criminal Law and Criminal Justice Administration. 10 paper presentations in various National and International seminars. Attended more than 10 FDP programs. Ph.D. in Law pursuing.



Avinash Kumar



Avinash Kumar has completed his Ph.D. in International Investment Law from the Dept. of Law & Governance, Central University of South Bihar. His research work is on "International Investment Agreement and State's right to regulate Foreign Investment." He qualified UGC-NET and has been selected for the prestigious ICSSR Doctoral Fellowship. He is an alumnus of the Faculty of Law, University of Delhi. Formerly he has been elected as Students Union President of Law Centre-1, University of Delhi. Moreover, he completed his LL.M. from the University of Delhi (2014-16), dissertation on "Cross-border Merger & Acquisition"; LL.B. from the University of Delhi (2011-14), and B.A. (Hons.) from Maharaja Agrasen College, University of Delhi. He has also obtained P.G. Diploma in IPR from the Indian Society of International Law, New Delhi. He has qualified UGC – NET examination and has been awarded ICSSR – Doctoral Fellowship. He has published six-plus articles and presented 9 plus papers in national and international seminars/conferences. He participated in several workshops on research methodology and teaching and learning.

ABOUT US

INTERNATIONAL JOURNAL FOR LEGAL RESEARCH & ANALYSIS
ISSN

2582-6433 is an Online Journal is Monthly, Peer Review, Academic Journal, Published online, that seeks to provide an interactive platform for the publication of Short Articles, Long Articles, Book Review, Case Comments, Research Papers, Essay in the field of Law & Multidisciplinary issue. Our aim is to upgrade the level of interaction and discourse about contemporary issues of law. We are eager to become a highly cited academic publication, through quality contributions from students, academics, professionals from the industry, the bar and the bench. INTERNATIONAL JOURNAL FOR LEGAL RESEARCH & ANALYSIS ISSN 2582-6433 welcomes contributions from all legal branches, as long as the work is original, unpublished and is in consonance with the submission guidelines.

SOCIAL MEDIA, FREE SPEECH, AND HATE SPEECH: LEGAL BOUNDARIES IN A POLARIZED SOCIETY

AUTHORED BY - RAKSHIT SHARMA

Abstract

In the digital era, social media has emerged as a transformative tool for communication, public engagement, and democratic participation. Platforms like Twitter, Facebook, and Instagram allow individuals to express opinions, share ideas, and mobilize for social causes on an unprecedented scale. However, the same digital platforms have increasingly become breeding grounds for hate speech, misinformation, and radicalization. Striking a balance between preserving free speech and limiting communication that promotes hate, bigotry, or public unrest is a difficult task.

This paper critically examines the legal and ethical boundaries of free speech and hate speech in the context of India's polarised society. It explores the constitutional guarantees under Article 19(1)(a) of the Indian Constitution and the permissible restrictions under Article 19(2). The analysis further includes key statutory provisions like Sections 124A, 153A, 295A, and 505 of the Indian Penal Code, and the regulatory framework under the Information Technology Act, 2000 and the IT Rules, 2021. Through judicial decisions such as *Shreya Singhal v. Union of India*, the paper evaluates how Indian courts have interpreted online speech.

Additionally, the study compares India's approach with global frameworks, especially the United States' First Amendment model and the European Union's regulations on hate speech and digital content. It examines how social media platforms enforce content moderation, their legal obligations, and the role of algorithms in amplifying harmful content.

The paper argues for a balanced legal regime that protects democratic expression while holding platforms and users accountable. It concludes with practical recommendations to ensure transparency, safeguard human rights, and maintain the integrity of online discourse. In doing so, the paper addresses one of the most pressing legal issues of our time: navigating free speech in an increasingly digital and divided society.

Introduction

In the past two decades, social media has transformed from a means of casual interaction to a central platform for public discourse. With platforms such as Twitter (now X), Facebook, Instagram, and YouTube reaching billions across the globe, digital spaces have evolved into modern-day public spheres where opinions are formed, social movements mobilized, and political narratives shaped. In India, with over 850 million internet users, social media holds the power to both empower and polarize, often acting as a parallel arena for civic engagement and dissent.¹

Freedom of speech and expression is one of the foundational principles of any democratic society. In India, this right is enshrined in Article 19(1)(a) of the Constitution and is central to the functioning of a vibrant democracy.² It enables individuals to question authority, advocate for change, and participate in democratic processes. The internet, and social media in particular, has expanded the reach of this right, giving even the most marginalized voices a platform to be heard.³

However, the digital age has also brought with it new challenges. The same platforms that allow for democratic dialogue have increasingly become hosts to hate speech, misinformation, targeted harassment, and radicalization.⁴ The anonymity and speed at which harmful content spreads online have raised serious concerns about public order, communal harmony, and the safety of individuals, especially minorities and vulnerable groups.⁵

This situation has triggered a demand for legal intervention to curb the misuse of online platforms. Yet, the regulation of digital speech poses a constitutional dilemma. Any attempt to curb harmful content must be balanced with the protection of fundamental rights. Over-regulation can lead to censorship and suppression of legitimate dissent, while under-regulation can result in unchecked harm.⁶

¹ Internet and Mobile Association of India, *Digital in India Report 2023*, available at <https://www.iamai.in/reports> (last visited June 15, 2025).

² INDIA CONST. art. 19(1)(a).

³ Justice A.P. Shah Committee Report on Privacy, *Government of India*, 2012.

⁴ Apar Gupta, "Hate Speech in the Digital Age," *Economic and Political Weekly*, Vol. 54, No. 30 (2019), pp. 18–22.

⁵ Nishith Desai Associates, *social media and Free Speech: Legal Landscape in India*, 2021, available at <https://www.nishithdesai.com> (last visited June 15, 2025).

⁶ Chinmayi Arun, "Hate Speech Laws in India," *Centre for Communication Governance Working Paper*, National Law University, Delhi, 2020.

This paper seeks to explore the legal boundaries of free speech and hate speech in the context of India's polarised digital landscape. It will examine the constitutional and statutory framework, judicial interpretations, the role of social media intermediaries, and global best practices. Through a socio-legal lens, the paper aims to assess whether India's current regulatory approach effectively balances democratic freedoms with the need to curb online harm, and what reforms may be required to navigate this evolving digital terrain.

Theoretical Foundations of Free Speech

Freedom of speech is the cornerstone of democratic societies, granting individuals the liberty to express opinions without undue interference. Philosophically, its roots trace back to classical thinkers like Socrates and were later expanded upon during the Enlightenment by thinkers such as John Locke and John Stuart Mill, who emphasized that open discourse is essential to individual liberty and the pursuit of truth.⁷

A central justification for free speech is the "marketplace of ideas" theory, which posits that all ideas, regardless of how controversial, should be freely expressed and tested through public debate. The best ideas, it is argued, will prevail over time.⁸ This concept, endorsed by Justice Oliver Wendell Holmes Jr., is vital for fostering democratic decision-making and an informed public.

Additionally, free speech is integral to individual autonomy.⁹ The ability to articulate one's thoughts is central to personal dignity, identity, and participation in public life. It empowers citizens to challenge authority, demand accountability, and engage in meaningful discourse.¹⁰ In this sense, speech is not merely a right but a means to preserve democratic structures.

However, the right is not absolute. Mill's "harm principle" suggests that expression can be restricted if it causes tangible harm to others.¹¹ Modern legal systems incorporate this by limiting speech that incites violence, spreads misinformation, or endangers public order. In culturally diverse societies, regulation becomes especially sensitive as it must protect

⁷ J.S. Mill, *On Liberty*, (1859), available at <https://www.gutenberg.org/ebooks/34901> (last visited June 15, 2025).

⁸ *Abrams v. United States*, 250 U.S. 616 (1919).

⁹ C. Edwin Baker, *Human Liberty and Freedom of Speech*, Oxford University Press (1989).

¹⁰ Owen M. Fiss, "Free Speech and Social Structure," 71 *Iowa L. Rev.* 1405 (1986).

¹¹ J.S. Mill, *supra* note 1, Chapter III.

vulnerable groups while preserving open discourse.¹² Thus, the legal and philosophical foundation of free speech reflects a dynamic balance between liberty and social responsibility.

Indian Constitutional and Legal Framework

The Indian Constitution guarantees the freedom of speech and expression as a fundamental right under Article 19(1)(a). This provision is central to the functioning of Indian democracy, enabling individuals to freely express opinions, disseminate information, and participate in public discourse.¹³ It encompasses various forms of expression, including spoken and written words, gestures, signs, art, and now extends to digital communication.¹⁴

However, this freedom is not absolute. Article 19(2) empowers the State to impose "reasonable restrictions" on this right 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, contempt of court, defamation, or incitement to an offence.¹⁵ The courts have interpreted "reasonable" to mean that restrictions must not be arbitrary, excessive, or disproportionate.¹⁶ This constitutional balance seeks to harmonise individual liberty with social order.

In addition to constitutional safeguards, several provisions under the Indian Penal Code (IPC) serve to regulate hate speech and prevent abuse of free expression:

- **Section 124A (Sedition):** Punishes speech or writing that incites hatred, contempt, or disaffection against the government.¹⁷ Though often criticized for being colonial in origin, it remains in force, despite growing demands for repeal or reform.¹⁸
- **Section 153A:** Penalizes the promotion of enmity between different groups on grounds such as religion, race, or language, and acts prejudicial to communal harmony.¹⁹ This section is frequently invoked in cases of inflammatory speeches and social media posts.
- **Section 295A:** Criminalizes deliberate and malicious acts intended to outrage religious feelings of any class by insulting religion or religious beliefs.²⁰ It is intended to

¹² Andrei Marmor, "What Is the Right to Free Speech?", *Philosophy & Public Affairs*, Vol. 43, No. 3 (2015), pp. 262–295.

¹³ INDIA CONST. art. 19(1)(a).

¹⁴ R. Rajagopal v. State of Tamil Nadu, (1994) 6 SCC 632.

¹⁵ INDIA CONST. art. 19(2).

¹⁶ Modern Dental College and Research Centre v. State of Madhya Pradesh, (2016) 7 SCC 353

¹⁷ Indian Penal Code, 1860, §124A

¹⁸ Gautam Bhatia, "The Case Against the Sedition Law," *The Wire*, 2020, available at <https://thewire.in/law/sedition-supreme-court> (last visited June 15, 2025).

¹⁹ Indian Penal Code, 1860, §153A.

²⁰ Indian Penal Code, 1860, §295A.

safeguard religious tolerance but has been criticized for its vague wording and potential misuse.

- **Section 505:** Covers the dissemination of statements or rumours that may cause fear, alarm, or incite offences against public tranquillity.²¹ It serves as a broad catch-all provision against provocative content, particularly in mass communication.

With the advent of the internet and the rise of social media platforms, traditional legal mechanisms have expanded to include cyber-specific regulations. The Information Technology Act, 2000 (IT Act) is the primary legislation governing online conduct in India. It was enacted to provide legal recognition to electronic communications and digital transactions. Over time, its scope has evolved to address content-related offences and intermediary liabilities.

The now-struck-down Section 66A of the IT Act, which criminalized sending offensive messages through communication service, was widely misused to arrest individuals for online posts, until it was declared unconstitutional in *Shreya Singhal v. Union of India*.²² This landmark decision upheld the primacy of free speech in the digital age while distinguishing between discussion, advocacy, and incitement.

To enhance regulatory control, the Government notified the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021. These rules impose due diligence obligations on intermediaries such as social media platforms, messaging services, and digital news outlets.²³ Key provisions include:

- Mandatory grievance redressal mechanisms
- Appointment of compliance officers for large platforms
- Time-bound takedown of unlawful content upon government or court direction
- Identification of the “first originator” of a message on encrypted services like WhatsApp in certain cases²⁴

While the IT Rules aim to curb misinformation and harmful content, they have raised concerns about privacy, surveillance, and censorship. Critics argue that they expand executive power without sufficient checks and threaten end-to-end encryption and press freedom.²⁵ Several

²¹ Indian Penal Code, 1860, §505.

²² *Shreya Singhal v. Union of India*, (2015) 5 SCC 1

²³ Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021.

²⁴ *Ibid.* Rule 4 (2)

²⁵ Internet Freedom Foundation, “Why IT Rules 2021 Threaten Online Freedom,” available at <https://internetfreedom.in> (last visited June 15, 2025).

petitions challenging the Rules are pending before constitutional courts.

In summary, India's legal framework seeks to maintain a balance between individual freedom and public interest, but faces growing challenges in regulating digital speech. The broad and sometimes overlapping language of IPC sections, combined with expansive IT rules, creates a legal ecosystem that can be both protective and repressive—depending on how it is interpreted and enforced. A coherent, rights-based approach is essential to ensure that the digital exercise of free speech aligns with constitutional values without undermining public safety or social harmony.

Judicial Interpretation and Key Cases

The Indian judiciary has played a critical role in interpreting the contours of free speech and determining the permissible limits on expression, particularly in the digital age. Through landmark decisions, courts have tried to reconcile the fundamental right to freedom of speech with the State's interest in maintaining public order, security, and communal harmony.

One of the most significant cases in this context is *Shreya Singhal v. Union of India*, where the Supreme Court struck down Section 66A of the Information Technology Act, 2000, holding it unconstitutional.²⁶ The provision criminalized the sending of “offensive” messages via electronic means, but its vague and overbroad language led to arbitrary arrests and suppression of online speech. The Court ruled that the section violated Article 19(1)(a) and failed the test of “reasonable restriction” under Article 19(2). It distinguished between discussion, advocacy, and incitement, and held that only speech that incites imminent violence or public disorder can be justifiably restricted.²⁷

In *Amitabh Thakur v. Union of India*, the Allahabad High Court examined whether provocative speech that disturbs public order should attract legal consequences.²⁸ The Court emphasized that while speech is a fundamental right, it is not absolute and must be exercised with a sense of responsibility. It upheld the idea that the State has a legitimate interest in preventing speech that could provoke violence, hatred, or communal unrest. The case reaffirmed that the threshold for restricting speech must be based on clear and present danger to public order, not on

²⁶ *Shreya Singhal v. Union of India*, (2015) 5 SCC 1.

²⁷ *Ibid.*, para 90.

²⁸ *Amitabh Thakur v. Union of India*, W.P. (CrI.) No. 20430 of 2012, decided on 3 July 2015 (All. HC).

subjective discomfort or political pressure.

Another noteworthy case is *Pravasi Bhalai Sangathan v. Union of India*, where the Supreme Court dealt with a petition seeking guidelines to curb hate speech.²⁹ While the Court acknowledged the harmful impact of hate speech, especially during elections and public rallies, it declined to frame guidelines, citing the existing legal framework. The Court held that it was the responsibility of the legislature to define hate speech more clearly and regulate it effectively. This decision reflected the judiciary's caution in venturing into legislative functions and highlighted the legal vacuum surrounding hate speech regulation in India.³⁰

Collectively, these cases illustrate the evolving judicial approach to balancing freedom of expression with societal interests. The Supreme Court and High Courts have reiterated that restrictions on speech must be narrowly tailored, and vague laws that can be misused to stifle dissent are incompatible with constitutional guarantees. They have also recognized the role of social media in amplifying speech, which makes the need for precision in law and proportionality in enforcement even more critical.

While the judiciary has laid down important principles, the inconsistency in application across courts, coupled with the absence of a statutory definition of hate speech, continues to pose challenges. As India's digital ecosystem expands, the role of courts in interpreting speech-related cases will remain crucial in preserving the balance between individual liberty and collective security.

Social Media Platforms and Content Moderation

Social media platforms such as Facebook, Twitter (now X), Instagram, and YouTube have fundamentally altered the dynamics of speech and information flow. They serve as modern public squares but operate under private control, creating a unique tension between user rights and corporate policies. The role and liability of these platforms—termed intermediaries under Indian law—have become central to discussions on digital governance and free speech regulation.

Under the Information Technology Act, 2000, intermediaries are granted conditional protection

²⁹ *Pravasi Bhalai Sangathan v. Union of India*, (2014) 11 SCC 477.

³⁰ *Ibid.*, paras 27–30.

through the “safe harbour” principle.³¹ Section 79 of the Act states that intermediaries shall not be held liable for user-generated content, provided they do not initiate, select, or modify such content and adhere to prescribed due diligence norms.³² This principle encourages platforms to host content without fear of constant litigation but also places a burden on them to remove unlawful content once notified.

Stricter compliance requirements for "significant social media intermediaries" (SSMIs) were imposed by the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, which widened the scope of intermediaries' legal responsibilities. A few examples of these measures are the establishment of grievance redressal officers, the release of monthly compliance reports, and systems to identify the original data source.³³ On the other hand, some think these rules could make state surveillance too easy and weaken end-to-end encryption.³⁴

Algorithmic amplification of hate speech has become an increasingly pressing issue in the past few years. Platform algorithms are designed to maximize user engagement, often by promoting sensational or polarizing content. This results in echo chambers, where users are repeatedly exposed to one-sided information, reinforcing biases and deepening ideological divides.³⁵ The lack of transparency in how algorithms function makes it difficult to assess accountability or implement regulatory checks.

Compounding this issue is the challenge of platform self-regulation. Most platforms maintain community guidelines that prohibit hate speech, harassment, and incitement. However, enforcement is often inconsistent and opaque. Content moderation practices rely heavily on automated systems, which may lack cultural sensitivity or contextual understanding—leading to wrongful takedowns or unchecked hate.³⁶ Moreover, political influence and economic incentives can skew moderation decisions, threatening the neutrality of these platforms.

³¹ Information Technology Act, 2000, § 79.

³² Ibid.

³³ Ibid. Rule 4.

³⁴ Internet Freedom Foundation, “Why the IT Rules, 2021 Are Concerning,” available at <https://internetfreedom.in/why-it-rules-are-dangerous> (last visited June 15, 2025).

³⁵ S. Tufekci, “Algorithmic Harms beyond Facebook and Google: Emergent Challenges of Computational Agency,” *Colorado Technology Law Journal*, Vol. 13 (2015), pp. 203–218.

³⁶ Tarleton Gillespie, *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*, Yale University Press (2018).

Another emerging concern is the use of dark patterns—manipulative interface designs that steer users toward certain behaviours without their full awareness.³⁷ These patterns can obscure privacy settings, encourage compulsive scrolling, or subtly guide users toward inflammatory content. Their use raises ethical and legal questions, particularly concerning informed consent and user autonomy in digital spaces.

In light of these challenges, there is growing advocacy for algorithmic accountability, transparency in content moderation, and greater regulatory oversight. While platforms play a crucial role in preserving online safety and civil discourse, their immense influence must be balanced with democratic values and user rights. The current Indian legal framework makes strides in this direction but requires further refinement to ensure that intermediaries act as responsible custodians of digital speech.

International Approaches and Comparative Analysis

The global legal landscape reflects varying approaches to regulating free speech and hate speech, shaped by each jurisdiction's historical, constitutional, and cultural context. While the United States adopts a near-absolute model of free speech, the European Union and countries like Germany impose stricter controls, especially against hate speech and misinformation. These models offer valuable insights for India as it navigates the complexities of digital speech regulation.

In the United States, the First Amendment of the Constitution guarantees broad protection of free speech. Courts have consistently held that even offensive or hateful speech is protected unless it incites "imminent lawless action."³⁸ The *Brandenburg v. Ohio* ruling established that only speech directly inciting violence can be curtailed.³⁹ This approach, while promoting robust discourse, often limits the government's ability to intervene against harmful online content, including hate speech.

In contrast, the European Union (EU) adopts a more restrictive framework. Article 11 of the EU Charter of Fundamental Rights protects freedom of expression but explicitly allows

³⁷ Arun Mohan Sukumar, "Dark Patterns and the Architecture of Online Manipulation," *The Hindu*, April 2022.

³⁸ U.S. CONST. amend. I.

³⁹ *Brandenburg v. Ohio*, 395 U.S. 444 (1969).

limitations to safeguard dignity, equality, and public safety.⁴⁰ The Digital Services Act (DSA) requires platforms to remove illegal content, including hate speech, swiftly and transparently.⁴¹ Additionally, the Audiovisual Media Services Directive (AVMSD) obligates member states to ensure online content does not incite hatred.⁴² These frameworks demonstrate how speech regulation can be embedded within human rights safeguards and due process mechanisms.

Germany's Network Enforcement Act (NetzDG) is a leading national example. Enacted in 2017, the law requires social media platforms with over two million users to remove "manifestly illegal content" (such as hate speech or incitement to violence) within 24 hours of receiving a complaint.⁴³ Non-compliance can result in hefty fines. Though effective in increasing compliance, the law has been criticized for encouraging over-removal of content due to strict deadlines.

India's regulatory efforts must strike a balance between protecting constitutional freedoms and curbing real-world harms. Borrowing from the EU's rights-based framework and Germany's enforcement model, while adapting them to Indian constitutional values, can offer a more democratic and effective path forward.

Political Polarisation and Online Radicalisation

In recent years, social media has increasingly become a breeding ground for political polarisation and radicalisation. While it has empowered citizens to engage in democratic discourse, it has also amplified divisive narratives, often fuelled by hate speech, misinformation, and partisan content. The rapid dissemination of such material not only distorts public debate but also contributes to the deepening of societal fault lines along political, religious, and caste-based identities.⁴⁴

Hate speech online plays a significant role in exacerbating social divisions. By reinforcing stereotypes and targeting specific communities, such speech fosters an "us vs them" mentality

⁴⁰ Charter of Fundamental Rights of the European Union, art. 11.

⁴¹ European Commission, *Digital Services Act*, available at <https://digital-strategy.ec.europa.eu> (last visited June 15, 2025).

⁴² Directive 2010/13/EU (AVMSD), as amended by Directive (EU) 2018/1808.

⁴³ *Netzwerkdurchsetzungsgesetz (NetzDG)*, Germany, 2017

⁴⁴ Sahana Udupa and Matti Pohjonen, "Digital Hate: The Global Conjunction of Extreme Speech," *social media + Society*, Vol. 5, No. 3 (2019).

that erodes mutual tolerance and democratic pluralism.⁴⁵ What might begin as isolated online rhetoric can, over time, normalize prejudice and even incite offline violence, especially in volatile socio-political environments.⁴⁶

The rise of troll armies, automated bots, and fake news factories has further distorted political narratives.⁴⁷ Coordinated campaigns are often used to discredit political opponents, manipulate public opinion, and flood social media feeds with emotionally charged content. These tactics can rapidly shift the tone of online discourse from civil disagreement to outright hostility. The anonymity of online platforms makes it easier for actors—domestic or foreign—to deploy misinformation and hate speech without accountability.⁴⁸

Another worrying phenomenon is the emergence of online mob justice and cancel culture. Individuals are increasingly “tried” in the court of public opinion based on viral posts, old comments, or controversial statements—often without context or legal recourse. While social media outrage can sometimes expose genuine wrongdoing, it also encourages knee-jerk reactions and character assassination, undermining due process and fostering a climate of fear and self-censorship.⁴⁹

Together, these trends signal a shift from digital participation to digital weaponization. The spread of hatred and disinformation not only threatens social harmony but also weakens the foundations of democratic engagement. Addressing this challenge requires a combination of legal reform, ethical platform design, algorithmic transparency, and above all, a collective commitment to responsible digital citizenship.⁵⁰

Challenges in Regulating Speech Online (with ILI-Style Footnotes)

Regulating speech in the digital age presents complex challenges, particularly in pluralistic democracies like India. While legal frameworks exist to curb harmful content, their effective

⁴⁵ UNESCO, *Countering Online Hate Speech*, 2015, available at <https://unesdoc.unesco.org/ark:/48223/pf0000233231> (last visited June 15, 2025).

⁴⁶ UN Special Rapporteur on Minority Issues, *Hate Speech and Social Media*, A/HRC/49/46 (2022).

⁴⁷ S. Woolley and P.N. Howard, *Computational Propaganda*, Oxford Internet Institute (2018).

⁴⁸ Mozilla Foundation, “YouTube Regrets: How the Platform’s Algorithm Amplifies Extremism,” available at <https://foundation.mozilla.org> (last visited June 15, 2025).

⁴⁹ R. Duffy and J. Ryan-Mosley, “Cancel Culture and Due Process in the Digital Age,” *MIT Technology Review*, 2021.

⁵⁰ Chinmayi Arun, “Extreme Speech in India,” in S. Udupa and M. Pohjonen (eds.), *Digital Hate: Comparative Perspectives*, Oxford University Press (2021).

implementation is often hindered by ambiguity, technological limitations, and the lack of institutional checks. One of the foremost issues is the vagueness surrounding the definition of “hate speech” and “public order.”⁵¹ Indian law does not provide a statutory definition of hate speech, leaving it open to broad interpretation.⁵² This ambiguity allows for subjective enforcement, where authorities may suppress dissent or criticism under the guise of maintaining public order.

Such vagueness creates a risk of over-censorship and leads to a chilling effect on free expression.⁵³ Individuals may self-censor out of fear of legal repercussions or social backlash, especially when political or religious sentiments are involved.⁵⁴ Laws that rely on undefined or overly broad terms can be weaponized against journalists, activists, and marginalized voices, undermining democratic discourse.

Another significant issue is algorithmic bias and opacity in content moderation. Social media platforms use artificial intelligence to filter and flag objectionable content. However, these algorithms are not neutral; they are trained on data sets that may reflect existing social or cultural prejudices.⁵⁵ As a result, minority viewpoints may be disproportionately flagged, while genuinely harmful content may evade detection due to contextual nuances that algorithms cannot grasp. The lack of transparency regarding how moderation systems function further complicates regulatory oversight.⁵⁶

Adding to the problem is the limited accountability of platforms and low levels of user awareness. While laws like the IT Rules, 2021 impose certain obligations on intermediaries, enforcement remains inconsistent.⁵⁷ Users often do not understand how to report content, appeal takedown decisions, or demand data transparency. Without digital literacy and clear grievance redressal systems, the average user is ill-equipped to navigate their rights in the online ecosystem.

⁵¹ Law Commission of India, *Report No. 267 on Hate Speech*, 2017.

⁵² *Ibid.*, p. 11.

⁵³ A.G. Noorani, “Free Speech and the Law,” *Frontline*, Vol. 31, No. 14 (2014).

⁵⁴ Apar Gupta, “The Chilling Effect of Ambiguous Laws,” *The Hindu*, March 2020.

⁵⁵ Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, St. Martin’s Press (2018).

⁵⁶ Algorithm Watch, “Automated Content Moderation: An Explainer,” available at <https://algorithmwatch.org> (last visited June 15, 2025).

⁵⁷ Internet Freedom Foundation, “Compliance and Criticism: One Year of IT Rules 2021,” available at <https://internetfreedom.in> (last visited June 15, 2025).

Recommendations and the Way Forward (with ILI-Style Footnotes)

The challenge of regulating digital speech lies not only in controlling harmful content but in doing so without infringing upon fundamental rights. To create a balanced legal ecosystem, India must adopt a multi-dimensional reform strategy that ensures accountability, transparency, and democratic legitimacy in the governance of online expression.

First, there is an urgent need for clear legislative definitions and thresholds for hate speech and related offences. Current laws such as Sections 153A, 295A, and 505 of the IPC are framed in broad and vague terms, making them susceptible to misuse.⁵⁸ A statutory definition of “hate speech” that draws from international human rights norms and focuses on incitement to violence, discrimination, or hostility can reduce subjectivity in enforcement.⁵⁹

Second, judicial oversight must be integrated into content takedown mechanisms. Presently, platforms are often required to remove content based on government orders, with limited scope for independent review.⁶⁰ A framework allowing courts or independent regulatory bodies to oversee such decisions would ensure proportionality and protect against arbitrary censorship.⁶¹

Third, public consultation should become a standard part of digital policymaking. Laws and rules affecting the digital sphere must not be framed unilaterally by the executive. Engaging stakeholders—including civil society, technologists, legal experts, and users—would enhance the legitimacy and effectiveness of such regulations.⁶²

Fourth, India must invest in digital literacy and ethical technology education. Citizens should be made aware not only of their rights but also their responsibilities online.⁶³ Public campaigns and curriculum reforms can instil values of critical thinking, empathy, and digital hygiene, which are essential in combating misinformation and online hate.

Finally, a sustainable solution requires balancing freedom with responsibility. Free speech is vital to democracy, but it must coexist with respect for diversity and human dignity. A rights-

⁵⁸ Law Commission of India, *Report No. 267 on Hate Speech*, 2017.

⁵⁹ UN Rabat Plan of Action on the prohibition of advocacy of national, racial or religious hatred, A/HRC/22/17/Add.4.

⁶⁰ Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021.

⁶¹ *Shreya Singhal v. Union of India*, (2015) 5 SCC 1.

⁶² Pratik Sinha, “Why Public Consultation Is Key in IT Governance,” *The Wire*, May 2021.

⁶³ Ministry of Electronics and Information Technology, *Digital Literacy Programme*, available at <https://meity.gov.in> (last visited June 15, 2025).

based approach, rooted in constitutional values and democratic principles, is necessary to ensure that India's digital future is both inclusive and secure.

Conclusion

The intersection of social media, free speech, and hate speech presents one of the most pressing legal and societal challenges in contemporary India. This paper has explored the constitutional and statutory framework governing speech, the interpretation by courts, and the role of intermediaries in shaping the digital discourse. It has highlighted how the lack of clear legal definitions, inconsistent enforcement, and opaque content moderation practices have created a regulatory vacuum—one that is easily exploited for political, social, or ideological purposes.

While India's Constitution guarantees freedom of speech under Article 19(1)(a), the permissible restrictions under Article 19(2) must be applied narrowly and proportionately.⁶⁴ Judicial precedents such as *Shreya Singhal v. Union of India* have reinforced the importance of protecting expression, especially in digital spaces, while also acknowledging the necessity of limiting speech that incites violence or disrupts public order.⁶⁵ However, the practical implementation of these safeguards often falters due to overbroad laws and unchecked executive powers.

There is an urgent need to shift towards a rights-respecting regulatory framework—one that is transparent, proportionate, and aligned with constitutional values. This includes enacting precise legislation on hate speech, enabling independent judicial oversight over takedowns, and ensuring due process for affected users.⁶⁶

Ultimately, the governance of digital speech cannot be the responsibility of the State alone. A truly democratic response requires a collaborative approach involving all stakeholders—government authorities, social media platforms, the judiciary, civil society, and users themselves.⁶⁷ Protecting freedom of expression while addressing digital harms is a shared responsibility, and only through inclusive dialogue, ethical technology, and legal reform can India navigate the complexities of its polarised digital society.

⁶⁴ INDIA CONST. arts. 19(1)(a), 19(2).

⁶⁵ *Shreya Singhal v. Union of India*, (2015) 5 SCC 1.

⁶⁶ Law Commission of India, *Report No. 267 on Hate Speech*, 2017.

⁶⁷ Chinmayi Arun, "The State of Content Regulation in India," *Centre for Communication Governance Working Paper*, 2020.