



REGULATING ONLINE HATE SPEECH IN INDIA: A CRITICAL ANALYSIS OF THE BNS, IT ACT AND INTERMEDIARY RULES

**Dr. Bhupender
Kumar Jodhta**

Professor & Principal, Awasthi College of Law, Nalagarh, Solan, Hp -174101

ABSTRACT

The expansion of social media has transformed how speech is created and circulated in India, strengthening democratic participation while also enabling the rapid spread of hate speech, communal hostility, and discriminatory content. India has no single, comprehensive statute defining "hate speech"; regulation is instead dispersed across constitutional provisions, the Bharatiya Nyaya Sanhita, 2023 (BNS), the Information Technology Act, 2000 (IT Act), and the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 (IT Rules), supplemented by judicial interpretation. Section 196 of the BNS criminalises the promotion of enmity through electronic communication, while Sections 79 and 69A of the IT Act establish intermediary safe harbour and blocking mechanisms respectively. The Supreme Court's decision in *Shreya Singhal v. Union of India* (2015) remains foundational in defining the constitutional limits of speech regulation under Articles 19(1)(a) and 19(2), while *Amish Devgan v. Union of India* (2021) situates hate speech within the framework of dignity, equality, and public order under Article 21. This paper critically examines the fragmented legal framework governing online hate speech in India, evaluates intermediary responsibility under the IT Rules, and analyses emerging challenges posed by algorithmic amplification, cross-border dissemination, anonymity, and AI-generated or synthetically generated content, particularly in light of the 2026 amendments to the IT Rules. It argues that a constitutionally sustainable regulatory model requires clear statutory standards, proportionality, procedural safeguards, and mechanisms responsive to technological change, so that genuinely harmful expression is restrained without producing a chilling effect on legitimate democratic discourse.

KEYWORDS : Hate Speech, BNS 2023, Information Technology Act, Intermediary Liability, Article 19(1)(a), Artificial Intelligence

1. INTRODUCTION

Article 19(1)(a) of the Constitution guarantees freedom of speech and expression, subject to reasonable restrictions under Article 19(2) on grounds such as public order, decency, security of the State, and incitement to an offence. Social media has expanded this freedom dramatically, allowing individuals to communicate with millions without the institutional gatekeeping traditionally associated with newspapers or broadcasting. Yet the same features that make online communication powerful — virality, algorithmic recommendation, and anonymity — also facilitate the rapid dissemination of hateful and discriminatory content. The central legal question is therefore not whether hateful expression should be restricted, but where the constitutional boundary lies between protected expression and punishable hate speech. Indian law addresses this through the BNS, the IT Act, the IT Rules, and constitutional jurisprudence, rather than through a single statute.

2. Conceptualising Hate Speech

No Indian statute offers a universally accepted definition of "hate speech." The term generally denotes communication that attacks, dehumanises, or promotes hostility against individuals or groups based on religion, race, caste, ethnicity, or language. Offensive, satirical, or provocative speech is not necessarily criminal; a constitutional democracy cannot treat every statement that causes discomfort as an offence. In *Amish Devgan v. Union of India* (2021), the Supreme Court examined hate speech through the lens of dignity, equality, and public order, treating it as capable of undermining the equal social status of targeted groups. Online, this distinction matters further because private platforms may remove content through moderation even where it falls short of a criminal offence, meaning platform moderation and criminal prohibition should not be conflated.

3. Constitutional Framework

Article 19(1)(a) protects controversial and unpopular opinion, not merely socially acceptable speech, but this protection is expressly subject to Article 19(2). In *Shreya Singhal v. Union of India* (2015), the Supreme Court struck down Section 66A of the IT Act as unconstitutionally vague and overbroad, holding that restrictions on online expression cannot be justified

merely by the possibility of misuse. Article 21, protecting life and personal liberty, has been read to include human dignity, producing a tension between the speaker's expressive freedom and the dignity and security interests of targeted communities. *Amish Devgan* situates this tension within a framework of constitutional balancing rather than the absolute priority of either speech or regulation.

4. Hate Speech Under the Bharatiya Nyaya Sanhita, 2023

Section 196 of the BNS criminalises the promotion, or attempted promotion, of disharmony or enmity between religious, racial, linguistic, or regional groups, and expressly recognises electronic communication, giving it direct relevance to online hate speech; Section 196(2) enhances punishment where the offence occurs in a place of worship or religious assembly. Section 197 addresses assertions and activities prejudicial to national integration, extending liability where expression contributes to hostility against particular groups. Other BNS provisions become relevant where content deliberately targets religious sentiment, though the mere invocation of religion does not automatically render a statement criminal; the statutory ingredients of the specific offence, including intent and context, must be established.

5. The Information Technology Act, 2000

The IT Act does not create a general offence of "online hate speech"; its significance lies in intermediary liability, safe harbour, and blocking powers. Section 79 grants intermediaries conditional protection from liability for third-party content, recognising that platforms cannot be treated as publishers of every user upload. In *Shreya Singhal*, the Court upheld Section 79 but read down Section 79(3)(b), holding that intermediary liability arises only upon actual knowledge through a court order or government notification concerning unlawful acts relating to Article 19(2). Section 69A authorises the Central Government, subject to procedural safeguards, to block public access to specified online information; the Court upheld this provision and the associated Blocking Rules, emphasising that such powers cannot become an unrestricted executive tool over online expression.

6. Intermediary Rules, 2021 and the 2026 Amendments

The IT Rules, 2021 impose due-diligence obligations on

intermediaries, requiring reasonable efforts to address content promoting enmity intended to incite violence, alongside grievance-redressal mechanisms. Significant social-media intermediaries face enhanced compliance obligations reflecting their greater systemic reach. In February 2026, amendments concerning synthetically generated information came into force, addressing deepfakes and AI-generated or manipulated content. This development is directly relevant to hate speech, since generative AI can now produce fabricated videos, manipulated audio, synthetic images, and multilingual hateful messages at scale, complicating questions of attribution, authenticity, and the identification of the responsible actor among creators, disseminators, platforms, and AI developers.

7. Judicial Approach

Shreya Singhal v. Union of India, (2015) 5 SCC 1, remains the leading authority: expression cannot be restricted merely because it is offensive, restrictions must fall squarely within Article 19(2), and intermediary liability must be read consistently with constitutional guarantees. *Amish Devgan v. Union of India*, (2021) 1 SCC 1, develops the dignity-and-equality dimension of hate speech regulation. *Pravasi Bhalai Sangathan v. Union of India*, (2014) 11 SCC 477, further considers the adequacy of existing legal mechanisms and the broader constitutional problem posed by speech capable of generating discrimination and hostility, informing ongoing debate on whether India needs a more comprehensive statute.

8. Intermediary Liability, the Chilling Effect, and Emerging Challenges

Two competing risks frame intermediary responsibility. Excessive liability may encourage over-removal of lawful speech to minimise legal exposure, producing a chilling effect on political criticism, satire, and academic debate. Insufficient responsibility, conversely, permits recommendation systems and engagement-based ranking to amplify harmful content regardless of its origin, raising the unresolved question of whether liability should extend to platform architecture and not merely the original speaker. Cross-border hosting, anonymous communication, and jurisdictional fragmentation further complicate investigation, requiring preservation of electronic evidence, subscriber disclosure, and cooperation with foreign providers, while legitimate anonymous political expression must be distinguished from anonymity used to commit offences.

9. Critical Analysis and Recommendations

India's framework has notable strengths: the BNS recognises electronic communication within enmity-related offences, the IT Act supplies an intermediary liability architecture, the IT Rules impose due-diligence obligations, and constitutional jurisprudence guards against arbitrary restriction. Its weaknesses include the absence of a unified statutory definition, fragmentation across criminal, technology, and constitutional law, inconsistent enforcement, an inadequate focus on algorithmic amplification, unresolved questions concerning AI-generated content, cross-border enforcement gaps, and the risk of chilling legitimate debate. A balanced model should adopt clear statutory standards distinguishing criticism from criminal hate speech, examine intent and context, apply proportionality between removal, blocking, and prosecution, provide procedural safeguards including notice and review, and require platform transparency regarding moderation and government requests. Specific reforms include clearer statutory terminology, specialised law-enforcement training on digital evidence and AI-generated content, human review of automated moderation in significant cases, stronger evidence-preservation mechanisms, and enhanced international cooperation for cross-border investigation.

Online hate speech sits at the intersection of constitutional liberty, criminal justice, and digital governance. India regulates the phenomenon not through a single statute but through the combined operation of the BNS, IT Act, IT Rules, and constitutional jurisprudence. *Shreya Singhal* establishes that restrictions on online expression cannot be vague or overbroad, while *Amish Devgan* underscores the relevance of dignity, equality, and public order. The emergence of AI-generated content, addressed partly through the 2026 IT Rules amendments, has added new complexity. A constitutionally sustainable model must protect both sides of the equation — the freedom to speak and the right of individuals and communities to participate in public life with dignity and equality — through standards that are clear, necessary, and proportionate.

REFERENCES

1. *Amish Devgan v. Union of India*, (2021) 1 SCC 1.
2. *Bharatiya Nyaya Sanhita*, 2023, No. 45, Acts of Parliament, 2023 (India).
3. Constitution of India, arts. 19, 21.
4. Information Technology Act, No. 21 of 2000, India Code (2000).
5. Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, as amended 2026 (India).
6. Ministry of Electronics and Information Technology, Government of India. (2026). Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026: Explanatory materials concerning synthetically generated information.
7. *Pravasi Bhalai Sangathan v. Union of India*, (2014) 11 SCC 477.
8. *Shreya Singhal v. Union of India*, (2015) 5 SCC 1.

10. CONCLUSION