
CONSTITUTIONAL SCRUTINY OF ONLINE HATE SPEECH IN INDIA: BALANCING FREE EXPRESSION AND STATE REGULATION

Vanlalhlupuui, B.Sc (Gargi College, University of Delhi), LL.B. (University of Delhi),
LL.M. (Sharda University)

ABSTRACT

With a focus on the digital realm, this essay explores how India's courts and constitutional framework have managed the conflict between protecting free speech and suppressing communication that incites hatred, animosity, or violence. It outlines the legal foundation (reasonable-restriction clauses and constitutional guarantees), statutory laws commonly used to combat hate speech, intermediate regulation and takedown procedures, and the Supreme Court's jurisprudential standards for when the State may impose speech restrictions. The study evaluates significant rulings and legislative actions that influence online content governance, including the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, *Shreya Singhal v. Union of India* (which overturned Section 66A), and previous precedents regarding the acceptable extent of restrictions. In order to create a fair, open, and ethical system for handling hate speech online in India, the paper's conclusion suggests doctrinal and policy changes.

Keywords: hate speech, freedom of speech

1. Introduction

The speed and reach of public communication have increased thanks to digital mediums. The Internet facilitated the quick dissemination of content that target groups based on religion, caste, race, gender, or geography as it emerged as the primary platform for political discussion, social criticism, and cross-cultural interaction. The two challenges that regulators and courts must deal with are (a) safeguarding the free flow of ideas that democratic engagement necessitates, and (b) prohibiting speech that jeopardizes public order, communal peace, or the safety and dignity of vulnerable groups. Article 19(1)(a) of the Constitution, which permits “reasonable restrictions,” a number of criminal provisions in the Indian Penal Code (IPC), and a statutory framework for intermediaries under the Information Technology Act and related regulations serve as the background against which these issues are litigated and regulated in India. The constitutional framework for assessing statutory and administrative action is provided by the judiciary’s rules on vagueness, proportionality, and the necessity of narrow tailoring, which have been crucial in defining the line between acceptable speech and criminal behaviour on the internet.¹

2. Concept of “Hate Speech” in Indian Law

There isn't a single legal definition of “hate speech” in India. Rather, expressions that instigate violence, inspire hatred, or offend religious emotions are prohibited by a number of criminal laws, and regulatory guidelines mandate that intermediaries respond to specific types of content. Operationally, “hate speech” is typically defined as communications that (i) deliberately or carelessly foster animosity between groups, (ii) amount to a malicious and intentional insult of religion, or (iii) incite or are likely to inspire violence or public disorder against a protected class. Instead of following a single doctrinal definition, courts employ fact-sensitive analysis since the content, intent, and context decide whether a given speech belongs within these manifestations.²

3. Constitutional Baseline: Article 19(1)(a) and the Reasonable-Restriction Framework

All people are guaranteed the right to free speech and expression under Article 19(1)(a),

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

² ADR India, “Details of Indian Penal Code: Sections 153A, 295A,” available at <https://adrindia.org>. Last visited on 9th December, 2025.

although Article 19(2) lists acceptable limitations in the interests of, among other things, state security, public order, decency or morality, defamation, and incitement to a crime. The Indian Supreme Court has consistently emphasized that limitations under Article 19(2) must be reasonable, explicit, and proportionate; ambiguous or unduly expansive clauses that grant authorities undue latitude to restrict expression have been overturned. *Romesh Thappar v. State of Madras*³ and other foundational rulings established the rule that censorship and preventive limitations must be strictly applied and cannot be used to stifle dissent or valid criticism. Later, more sophisticated proportionality analysis emerged; the court mandated a strong relationship between the goal of the restriction and the methods employed to accomplish it.⁴

The following are important aspects of the constitutional criteria that affect the regulation of hate speech online:

1. Legitimate goal: According to Article 19(2), the restriction must aim to achieve a recognized goal, such as upholding public order or preventing incitement to violence.⁵
2. Legality: Vague criminal provisions are unconstitutional; any restriction must be supported by a law (a statute or rule) that is predictable in its application.⁶
3. Proportionality and Necessity: The action must be both proportionate to the harm it aims to avert and necessary to accomplish the justifiable goal. Courts frequently weigh the danger of expression against the social cost of suppression.⁷

These restrictions indicate that general takedowns or arrest powers without explicit guidelines or procedural protections will come under judicial scrutiny when it comes to online speech.

4. Statutory and Regulatory Architecture Governing Online Speech

4.1 Indian Penal Code provisions frequently used against hate speech

Despite the fact that the IPC predates the internet age, a number of its provisions are frequently used to regulate online speech that is thought to incite conflict among communities or offend

³ *Romesh Thappar v. State of Madras*, AIR 1950 SC 124 (India).

⁴ *Shreya Singhal*, supra note 1.

⁵ *Ibid*

⁶ Indian Kanoon, Constitutional commentary on Article 19(2), <https://indiankanoon.org>. Last visited on 9th December, 2025.

⁷ *S. Rangarajan v. P. Jagjivan Ram*, (1989) (India) (2) SCC 574.

religious sentiments:

- Section 153A (Promoting enmity between different groups): Prohibits acts that foster animosity, hostility, or animosity between classes on the basis of caste, religion, race, place of birth, domicile, or language.⁸
- Section 295A (Deliberate and malicious acts intended to outrage religious feelings): Penalizes statements or deeds that are purposefully meant to disparage a class's religion or religious beliefs. According to court interpretations, this clause necessitates wilful and malevolent purpose.⁹
- Section 505 (Statements conduced to public mischief): Makes it illegal to make statements that cause fear or panic, including those that could lead someone to commit a crime against the state or disturb the peace.¹⁰

The IPC's provisions apply to online communications in the same way as they do to offline speech; nevertheless, its adaptation to the Internet presents unique issues related to scale (viral re-sharing), authorship (anonymous accounts and intermediaries), and context (political or satire speech). As a result, courts have emphasized that speech does not always fall under the purview of these penal prohibitions just because it is objectionable or may cause wounded feelings; intention and connection to public disorder continue to be crucial.¹¹

4.2 Information Technology Act, intermediary liability, and the 2021 Rules

The primary legislative framework for digital governance in India is the Information Technology Act, 2000 (IT Act). The control of hate speech on the internet revolves around two key elements:

- Intermediary safe harbour: Intermediaries (platforms, ISPs) are exempt from liability for third-party content under the IT Act and its regulations as long as they adhere to due

⁸ Supra note 2.

⁹ Ibid.

¹⁰ Copyright.lawmatters.in, Section 505 commentary, <https://copyright.lawmatters.in>. Last visited on 9th December,2025.

¹¹ iPleaders, “Application of IPC to Online Speech,” <https://blog.ipleaders.in>. Last visited on 9th December,2025.

diligence standards and remove or terminate access to unlawful content upon acquiring actual knowledge, by means of a court order or a proper governmental notice.¹²

- Intermediary Guidelines and Digital Media Ethics Code Rules, 2021: The 2021 Rules revised the 2011 Intermediary Guidelines and added stricter due-diligence requirements for “significant social media intermediaries,” such as time-bound takedown requirements for content identified by the government or by irate users, traceability requirements for specific messaging services, and grievance redressal procedures.¹³

4.3 Content blocking and emergency powers

In addition to takedowns, the IT Act grants the administration particular authority to restrict access to content for reasons of public order, sovereignty, or to stop encouragement to criminal activity. For example, Section 69A allows for the barring of public access to information under specified procedures; courts must provide procedural protections and a level of rational decision-making for such blockages.¹⁴

5. Leading Judicial Authority and Doctrinal Tests

5.1 *Shreya Singhal v. Union of India (2015)* and the invalidation of Section 66A

For internet expression in India, the Supreme Court's finding in *Shreya Singhal v. Union of India* is a landmark decision. The Court declared that Section 66A of the IT Act breached Article 19(1)(a) due to its overbreadth and vagueness, as well as its failure to meet the Article 19(2) test of reasonable restrictions.¹⁵

For the purpose of regulating hate speech online, *Shreya Singhal v. Union of India* has three enduring doctrinal messages:

- Vagueness doctrine: Criminal prohibitions must be adequately explicit; ambiguous

¹² PRS Legislative Research, “IT (Intermediary Guidelines) Rules, 2021,” <https://prsindia.org>. Last visited on 9th December, 2025.

¹³ MeitY, Ministry of Electronics & IT, “Intermediary Guidelines and Digital Media Ethics Code Rules, 2021,” <https://www.meity.gov.in>. Last visited on 9th December, 2025.

¹⁴ TIME, “India Strikes Down Controversial Law Banning ‘Offensive’ Online Content,” <https://time.com>. Last visited on 9th December, 2025.

¹⁵ *Shreya Singhal*, supra note 1.

terms such as “offensive” are unconstitutional.¹⁶

- Proportionality and least-restrictive means: When tailored, less invasive remedies can achieve justifiable goals, the State cannot impose broad criminal sanctions.¹⁷
- Procedural safeguards for intermediaries and users: Targeted criminal sanctions cannot be imposed by the state; instead, platform shutdown and blocking must be limited by clear guidelines and channels for redress.¹⁸

5.2 Proportionality and “clear and present danger”: *S. Rangarajan* and other precedents

The Supreme Court’s ruling in *S. Rangarajan v. P. Jagjivan Ram* emphasized that restrictions are only acceptable in situations when there is a genuine and immediate threat, and that freedom of expression cannot be stifled by hypothetical or distant worries of retaliation.¹⁹

The Rangarajan test, which requires the State to demonstrate a close connection between the expression and the threat being prevented, is frequently used in conjunction with proportionality analysis. Courts and regulators must evaluate the context, the speaker's goal, the audience's expected response, and the immediate danger of harm when it comes to online hate speech.

5.3 *Pravasi Bhalai Sangathan* (2014) and judicial approaches to hate speech

The Supreme Court addressed political and communal discourse that endangered societal harmony in *Pravasi Bhalai Sangathan v. Union of India* (2014).²⁰ The Court acknowledged that hate speech might create obstacles to historically marginalized groups’ full participation in democratic society by drawing on similar precedents, such as *Saskatchewan v. Whatcott in Canada*. The directives and guidelines that emerged from *Pravasi Bhalai Sangathan v. Union of India* emphasized the necessity of suitable procedures to deal with hate speech while also warning against excessive criminalization that would stifle reasonable discussion. The ruling

¹⁶ Ibid

¹⁷ Global Freedom of Expression, online resources, <https://globalfreedomofexpression.columbia.edu/cases/bhasin-v-union-of-india>. Last visited on 9th December, 2025.

¹⁸ Supra note 12.

¹⁹ S. Rangarajan, supra note 7.

²⁰ *Pravasi Bhalai Sangathan v. Union of India*, Writ Petition (Civil) No. 157/2013, (2014) 11 SCC 477, ILDC 2827 (IN 2014), 12th March 2014, India.

emphasized the necessity for procedural justice and the protection of dissent, but it also supported legislative and administrative action.

5.4 Privacy, information and digital contexts: *K.S. Puttaswamy (2017)* and implications for online speech

The concepts of legality, need, and proportionality found in *K.S. Puttaswamy v. Union of India* are useful for regulating digital speech, even if the case is mainly about privacy.²¹ Courts have used the nine-judge bench's three-part test (existence of law, legitimate state aim, and proportionality in means) to assess data surveillance, intermediary obligations, and restrictions on digital expression, all of which have an impact on how hate speech is detected, looked into, and addressed online. Thus, the privacy finding enhances the constitutional toolkit for examining government authority over digital material and data stored on platforms.

6. Evaluating State Regulation: Strengths, Pitfalls and Operational Challenges

6.1 Strengths

- Multi-pronged toolbox: India has a number of legal tools that allow for targeted action against speech that crosses the line into incitement or planned violence, including criminal law (IPC sections), the IT Act and related Rules, and criminal procedure rules. These can be useful in reducing real risks to public order when implemented with proper process.
- Judicial guardrails: Strong constitutional protections against ambiguous or disproportionate limitations are provided by landmark rulings such *Shreya Singhal v. Union of India*, *S. Rangarajan v. P. Jagjivan Ram*, and *Justice K.S. Puttaswamy (Retd.) v. Union of India*, which put the onus on the State to defend speech violations. These precedents aid in preventing the arbitrary repression of minority opinions and dissent.
- Hybrid constitutional-regulatory model: A blended framework that seeks to prevent damaging speech while upholding democratic norms and freedom of expression is created by the combination of statutory requirements, executive regulations, and

²¹ *Justice K.S. Puttaswamy (Retd.) v. Union of India*, (2017) 10 SCC 1 (India).

constitutional jurisprudence.

- Structured accountability under the 2021 Rules: The 2021 Intermediary Guidelines and Digital Media Ethics Code Rules establish explicit procedural obligations for social media intermediaries, such as the need for timely takedowns, grievance officers, and response deadlines. These responsibilities improve India's ability to quickly remove dangerous content, especially when it goes viral.
- Intermediary safe-harbour system: Digital platforms are immune from third-party material under the IT Act, but only if they adhere to the required due diligence requirements. This strategy achieves a balance by encouraging platforms to present a range of opinions while still holding them accountable for taking action against dangerous or illegal information when they are alerted.
- Availability of content-blocking mechanisms: One of the government's most organized instruments for limiting access to dangerous digital content is Section 69A of the Information Technology Act. This provision is backed by a thorough, rule-based process outlined in the Information Technology (Procedure and Safeguards for Blocking for Access of Information by Public) Rules, 2009, in contrast to ad hoc or arbitrary takedown demands. In order to improve openness and legitimacy, the procedure calls for an evaluation of possible threats, a review by a designated committee, and the recording of written, well-reasoned directives. In times of emergency, such as communal unrest or crises driven by misinformation, this framework allows the State to react quickly to content that could jeopardize national security, disturb public order, or instigate violence. The method is further strengthened by the existence of emergency blocking capabilities, which are subject to future review. This ensures that bad material can be stopped even before it causes harm in the real world.

6.2 Pitfalls and operational challenges

- Vagueness and prosecutorial discretion: A number of the IPC's anti-hate speech provisions were created before the digital age; their wording can be ambiguous and occasionally interpreted in ways that include obnoxious but non-violent speech. This results in criminal complaints and arbitrary arrests, which stifle free speech. Although

the *Shreya Singhal v. Union of India* recognized and particularly addressed this issue with regard to Section 66A, there are still uncertainties in IPC jurisprudence.

- Speed vs. due process in takedowns: Content can be quickly removed thanks to executive takedown/blocking capabilities and intermediary rules. While speed is crucial to stop offensive content from going viral, hasty takedowns without clear guidelines or significant compensation could stifle legitimate expression and weaken impacted speakers. Critics contend that although the 2021 Rules raised procedural requirements for platforms, they actually do not sufficiently safeguard user rights.
- Scale and context: Platforms host vast amounts of content; it takes careful evaluation to discern between speech that poses a genuine risk of violence and speech that is abusive, satirical, or critical. Over-reliance on keyword screening and automated moderation can result in false positives, while under-moderation can allow harm. Courts have advised against substituting a precise determination of intent and substantial injury with criminal law.
- Traceability vs. privacy: The capacity to identify speakers is enhanced by certain regulations (and law enforcement requirements) for “traceability” of message originators, which aids in the implementation of laws against incitement. However, if users are afraid of intrusive surveillance, these tracking methods could stifle communication and pose privacy problems of their own. According to the *Justice K.S. Puttaswamy (Retd.) v. Union of India* privacy doctrine, such actions must be evaluated for need and proportionality.
- Political economy of enforcement: Political influences or selective policing may be reflected in enforcement practices; prosecutions or preventive deletions may be less rigorously pursued in cases where politically influential actors spread hate speech.

7. Comparative insights to Indian practice

Other democracies' legislative and judicial actions provide instructive differences. For instance, some jurisdictions combine strong notice-and-takedown procedures, judicial oversight of blocking orders, and specific criminal laws requiring purpose and likelihood of impending violence. When debating how to balance freedom of expression with collective

dignity, the Supreme Court in *Pravasi Bhalai Sangathan v. Union of India* itself referred to Canadian doctrine in *Whatcott* case. Comparative experience indicates that the best practices for democratic states looking to restrict hate speech without stifling legitimate debate include clear statutory language, procedural protections for challenged speakers, independent review of takedown/blocking decisions, and targeted remedies calibrated to risk.

8. Recommendations: A rights-aligned approach to online hate speech

The following suggestions, which are based on operational realities and theological concepts, seek to balance the need to prevent group-targeted harm online with the virtues of free expression:

- Legislative clarity and mens rea requirements: Modify or modify current penal statutes such that before criminal liability for online communications attaches, there must be a demonstrable mens rea (purpose or knowledge) and a proximate risk of public disorder or violence. Criminal penalties shouldn't be based on ambiguous terminology like "offensive." Legislative drafting should be guided by the criterion established in *Shreya Singhal v. Union of India*, which emphasizes clarity and limited tailoring.
- Statutory guidance for platform takedowns: The intermediate system should mandate transparent notification procedures, justified takedown orders, and an unbiased appeals process for content creators to seek timely review. Time-bound government takedown orders must be accompanied by explanations and subject to judicial or quasi-judicial scrutiny in order to reduce arbitrariness. Even though the 2021 Rules move in this direction, they may be strengthened by procedural safeguards and independent oversight.
- Proportional civil remedies and counter speech incentives: Non-criminal solutions such as de-amplification, demotion, labelling, corrections, and civil accountability should be used for verifiable harms, whereas criminal solutions must be avoided as much as possible. State policy should encourage platforms to prioritize de-escalation, context preservation, and counter speech interventions above immediate removal when doing so would limit free speech.
- Privacy-sensitive traceability standards: Traceability requirements for message

originators must pass a proportionality test and be restricted to significant national security or public order situations. The privacy rules established in *Justice K.S. Puttaswamy (Retd.) v. Union of India* would be violated by a broad, comprehensive tracing of everyday political speech. Any mandate requiring intermediaries to divulge identifying information need to be carefully focused and approved by the courts.

- Capacity building and evidence standards for law enforcement: Training in digital evidence, context analysis, and legal standards for incitement and public order could support police investigations into hate speech on the internet. To avoid abuse, a higher standard of proof should be necessary prior to an arrest or criminal prosecution.
- Independent oversight and transparency reporting: Periodically, platforms and their executives should release transparency reports that include information on removal requests (both private and governmental), the grounds used, and the platform's reaction. A judiciary-backed tribunal or independent regulator might examine widespread takedown trends and guard against biased enforcement.
- Public education and counter-narratives: To make the public less vulnerable to manipulation and inflammatory propaganda, democracies must engage in media literacy and community conversations that strengthen resistance against hate narratives. The socioeconomic causes of intergroup animosity cannot be eliminated by legal regulations alone. The Court's concerns about participation impediments in *Pravasi Bhalai Sangathan v. Union of India* and social science findings are the basis for this recommendation.

9. Conclusion

India's constitutional assessment of hate speech on the internet runs at the nexus of strong free-speech protection with the justifiable need to uphold social peace, public order, and human dignity. Important safeguards have been provided by Supreme Court jurisprudence, particularly *Shreya Singhal v. Union of India*, *S. Rangarajan v. P. Jagjivan Ram*, *Pravasi Bhalai Sangathan v. Union of India*, and the privacy doctrine in *Justice K.S. Puttaswamy (Retd.) v. Union of India*, laws must be precise, targeted, and proportionate; procedural safeguards must prevent arbitrary enforcement; and digital measures must respect privacy and the rule of law.

Developing a system that prevents and corrects harm to groups while maintaining the plurality and contestation necessary for democratic life is the regulatory issue of the future. This calls for judicial supervision, platform responsibility, legal precision, and public investment in anti-hate resistance. Combining these components will allow India to combat online hate speech in a way that is both constitutionally sound and balanced, safeguarding both the free flow of ideas and the dignity of marginalized communities.

REFERENCES

Primary legal materials

1. *Shreya Singhal v. Union of India*, (2015) 5 SCC 1 (India).
2. *Romesh Thappar v. State of Madras*, AIR 1950 SC 124.
3. *S. Rangarajan v. P. Jagjivan Ram*, (1989).
4. *Pravasi Bhalai Sangathan v. Union of India*, Writ Petition (Civil) No. 157/2013, (2014) 11 SCC 477, ILDC 2827 (IN 2014), 12th March 2014, India.
5. *Justice K.S. Puttaswamy (Retd.) v. Union of India*, (2017) 10 SCC 1

Statutes and Rules

1. Information Technology Act, 2000 (as amended) - provisions on intermediary safe-harbour, blocking and investigative powers. (Government of India)
2. Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021
3. Indian Penal Code, 1860 - Sections 153A, 295A, 505 (and related provisions criminalising promotion of enmity and outrage to religion).

Select secondary sources

1. PRS Legislative Research, “The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021” - explanatory note and analysis.
2. PRS Legislative Research
3. Time Magazine, “India Strikes Down Controversial Law Banning ‘Offensive’ Online Content” (overview of *Shreya Singhal v. Union of India* decision and its significance).
4. Academic and NGO commentaries on hate speech provisions and jurisprudence (e.g., analyses of Sections 153A/295A, reports accessible through NGO/academic portals).