# PROTECTING FREE SPEECH IN THE DIGITAL AGE: CONSTITUTIONAL CHALLENGES OF ONLINE REGULATION AND HATE SPEECH LAWS

Khushboo Rupani, Research Scholar (Law), Vikrant University, Gwalior

Dr. Vir Narayan, Dean (Legal School), College: Vikrant University, Gwalior

Mr. Prashant Rao Mulik, Assistant Professor (Legal School), Vikrant University, Gwalior

## ABSTRACT

The article focuses on the constitutional issues faced by online speech regulations in India and suggests legal and policy reforms that will not only safeguard the essence of Article 19(1)(a) but also mitigate the impact of hate speech, misinformation, and the use of synthetic media technologies. The study also examines the existing legal and regulatory frameworks, especially the Information Technology Act of 2000 (particularly the controversial Sections 69A and 79), the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules of 2021, and the overlapping punishment under the Bharatiya Nyaya Sanhita, such as Sections 194, 297, and 351, demonstrating how the three factors of vagueness, overbreadth, and fragmented enforcement are contributing to chilling effects on political dissent, satire, and legitimate criticism. The study analyzes online regulation through the lens of the constitutional criteria established by Article 19(2) (legitimate aim, proximate nexus, proportionality) and related constitutional values such as privacy (Article 21), and posits that the lack of technical clarity on mens rea and proximate harm leads to the possibility of misuse. The comparison of lessons from the United States (Section 230 and First Amendment jurisprudence) and the European Union (DSA-style duties and rights-based balancing) directs domestic options: maintain intermediary immunity under conditions but change discretionary rules into statutory, time-limited due-process protections (notice, counter-notice, independent appeals), restrict criminal jurisdiction by intent-based thresholds, and introduce targeted, technology-specific offences for malicious uses of deepfakes and synthetic media. Moreover, the article underlines execution: requirement for capacity building for forensic investigation, anti-SLAPP remedies, mandatory transparency reporting, and judicial safeguards for traceability and data access. The proposed framework is intended to reinforce the area of platform governance, policing practice, and individual rights in

such a way that the regulation only suppresses the truly dangerous speech and at the same time India's digital age enjoys a strong public sphere.

**Keywords:** Free speech; Information Technology Act; hate speech; intermediary liability; privacy (Article 21).

## 1. INTRODUCTION

The digital realm brings along the need to rethink the safeguarding of free speech in India. The main reason for this transformation is the fact that online channels, algorithmic amplification, and global distribution in real-time are actually converting private utterances to public happenings. According to Article 19(1)(a) of the Indian Constitution[1], everyone has the right to express their opinions and thoughts freely. In contrast, the Constitution allows the government to impose "reasonable restrictions" based on specific grounds like national security, public order, and morality per Article 19(2). These constitutional provisions are the foundation for judging any law or rule that controls digital communication. (Constitution of India, 1950).

To control online communication, the main methods that are used today are statutory law and administration rules. The Information Technology Act, 2000 grants powers, among others, the provisions of Section 69A (the blocking of public access to information) and Section 79[2] (conditional liability exemption for intermediaries). The use of these provisions, when viewed along with the administrative steps, decides the ways and the times when the State or private platforms can block or take down certain content. In addition, the very recent Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 impose even more duties on the part of the intermediaries: provision of grievance handling mechanisms, requirements of traceability, and compliance duties that significantly influence the behavior of the platforms and the privacy of the users. Directly related to these rules are criminal statutes that are often enacted in the context of online activity like the Bharatiya Nyaya Sanhita, which has Section 194 (causing conflict), Section 297 (deliberate and vicious acts intended to provoke religious feelings) and Section 351 (statements that may lead to public disturbance) among others. They are frequently brought into the online arena,

[1] INDIA CONST. art. 19, cl. 1, sub-cl. (a).
[2] Information Technology Act, No. 21 of 2000

creating concerns about expanse and misuse when applied to short-lived postings on social media. (Ministry of Electronics & Information Technology, 2021).

The interaction of these legal instruments at various levels generates doctrinal and practical issues: The risks to the political and artistic freedom of expression from the vague legislative wording and extensive criminal laws in the case of the latter are much too great; the obligation to intermediaries may eventually shift the content policing role from the courts to the private companies (which will be very reluctant to bear it and will thus remove more content than is necessary to avoid the risk of being sued); and the requirement of traceability may lead to conflict with encryption and privacy, thus becoming a question of Article 21 values. The same time without losing privacy and other rights, the proper decisions cannot be made if the new technologies like deepfakes, synthetic media, and automated content-moderation system are not taken into account as they create evidentiary and attributional difficulties. Thus, the policymaking process that involves the use of the proposed restrictions should be very careful so as not to interfere with the very core of Article 19(1)(a); these restrictions should be very narrowly tailored and the existing constitutional tests of legitimate aim, proximate nexus, and proportionality should be applied; reforms should be accompanied by the clarification of mens rea for online offences, proofing of statutory language to eliminate arbitrary enforcement, and establishment of transparent independent review mechanisms for taking down and blocking decisions. The contemporary debates on policies and the proposals for changes in IT legislation given by the government also suggest that the government is now placing more emphasis on the AI-mediated harms and the accountability of the platforms while at the same time the need for doctrinal refinement that brings statutory powers in line with constitutional safeguards is becoming more and more urgent.[3]

### 1.1 Objectives of the Study

The objectives of the research are as follows:

1. To scrutinize the constitutional boundaries of Article 19(1)(a) in the digital

[3] PRS Legislative Research, The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 — PRS Analysis (2021)

age;

2. To investigate legal statements—mainly Sections 69A and 79 of the Information Technology Act, 2000—and the IT (Intermediary Guidelines) Rules, 2021;

3. To evaluate the applicability and the potential of misuse of BNS provisions like 194, 297, and 351 on the internet;

4. To detect doctrinal deficiencies leading to chilling effects on free expression; and

5. To put forward legal and procedural reforms focusing on precision to ensure free speech and, at the same time, control online harms.

### 1.2 Research Methodology

The research in this paper is conducted through a doctrinal method of study which basically uses statutory texts, constitutional provisions, government rules, and authoritative secondary literature to extract the legal principles and interpretive directions. The main materials consist of the Constitution, the Information Technology Act, 2000 (particularly Sections 69A and 79), the **Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021**[4], and relevant provisions of the Bharatiya Nyaya Sanhita like Sections 194, 297, and 351. Scholarly articles, legislative histories, government policy notes, and comparative statutory materials are among the secondary sources. The analysis comprises the systematic interpretation of statutory language, legislative intent, and doctrinal tests (legitimate aim, proximate nexus, proportionality, vagueness), inter-statutory evaluation, and the discovery of normative deficiencies leading to overbreadth or arbitrariness. Wherever relevant, comparative references to regulatory models are used to support reform proposals. Doctrinal method limitations—e.g. restricted empirical assessment of enforcement practices—are recognized, and recommendations point to the necessity of complementary empirical research in certain areas.

[4] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, G.S.R. 139(E) (India)

## 2. THE CONSTITUTIONAL FRAMEWORK THAT SUPPORTS FREE SPEECH IN INDIA

The protection of free speech in India is made possible by and relies on the Constitution's Article 19(1)(a), which provides all citizens the right to freedom of speech and expression. This is a far-reaching guarantee because the Constitution is not only looking at the traditional media, but it is also covering new media, such as digital and online communications, thus enabling speech on the internet or social media platforms to be included. The very inclusion of this area is nothing but a reaffirmation of the democratic aspiration for open dialogue and public discussion in a society consisting of various intermingling cultures.[5]

Notwithstanding, this right does not apply in every instance. “Reasonable restrictions” in the interest of the sovereignty and integrity of India, public order, decency or morality, friendly relations with foreign countries, contempt of court, defamation, or incitement to an offence, among other things, are the statutory grounds created by Article 19(2). The effect of the constitutional provision is such that it establishes a balance between the two extremes of freedom of expression and legitimate restrictions imposed by the State. (Srivastava, S, 2023)

To add, free speech must be placed in the context of the other constitutional values, such as dignity, equality, and privacy, apart from Articles 19(1)(a) and 19(2). In the digital age, where speech is allowed through private platforms and is often linked to user data, privacy is of great importance. The right to informational privacy and personal dignity are the deciding factors in what is regarded as acceptable regulation (for instance, traceability, data-retention, notice-and-takedown), meaning henceforth that any limitation placed on speech is to be weighed not only against public-order concerns but also against individual autonomy and equality. (Singh, D. P., 2025)

The constitutional framework for free speech in India consists of a fundamental right to freedom of speech and expression under Article 19(1)(a), certain limitations as per Article 19(2) and wider constitutional values like dignity, privacy, and equality that are more

[5] D.P. Singh, An Analysis of the Article 19(1)(a) and Article 19(2) of the Indian Constitution and Distorting Form of Freedom of Speech and Expression in the Era of Social Media in India, SSRN Scholarly Paper No. 5100601 (2025)

pronounced in the networked, digital world — all of which become very relevant in the networked, digital era.

## 3. STATUTORY & REGULATORY ARCHITECTURE GOVERNING ONLINE SPEECH

### 3.1 Information Technology Act, 2000 — major provisions and evolution (Sections 66A, 69A, 79 etc.)

The Information Technology Act, 2000 (IT Act, 2000) is at the top of the legislative pyramid in India that governs online and digital speech. The Act was enacted to combat cybercrimes, facilitate electronic commerce, and streamline electronic governance. The IT Act established the basic rules for the regulation of intermediaries and the government's online content control.

The IT Act empowered the central government under Section 69A to issue directions to the organizations or intermediaries to deny public access to internet information if such action is required for the protection of sovereignty or territorial integrity of India, state security, maintaining friendly relations with other countries, ensuring public order, or preventing a non-bailable offence from being committed. On the other side, Section 79 grants a "safe harbour" to intermediaries, which means they are not liable for the contents posted by others as long as they do their part and remove the "unlawful information" as soon as they have actual knowledge, or get a court or government order[6] (Wadhwa, S., 2025).

### 3.2 The Intermediary Guidelines and Digital Media Ethics Code Rules, 2021 — responsibilities, fair trial rights and notification/traceability requirements

The Intermediary Guidelines and Digital Media Ethics Code Rules, 2021 ("IT Rules, 2021") were published as part of a regulatory regime for online intermediaries and digital media that the IT Act allowed under its rule-making power (Section 87). The Rules require intermediaries, among others, to put in place grievance redressal mechanisms, appoint grievance officers, subject themselves to

[6] Wadhwa, S. (2025). Content blocking orders and status of digital rights: Assessment of two key verdicts in India. Information & Communications Technology Law, 34(1), 44-61

takedown notices, keep records, save data for a certain time, and above all make sure that originators are traceable (especially for "significant social media intermediaries").

### 3.3 Criminal law backstops: relevant provisions of the BNS (194, 297/98, 351, sedition provisions)

At the same time, the application of criminal laws to online speech remains in effect. Through provisions of the Bharatiya Nyaya Sanhita (BNS), such as Sections 194 (promoting enmity between groups), 297 (deliberate and malicious acts intended to outrage religious feelings), 351 (statements conducive to public mischief), sedition laws (e.g., formerly Section 150), and other penal provisions relating to defamation and incitement to offence, are frequently invoked in the digital context. These criminal provisions act as "backstops," allowing the State to prosecute individuals for speech — including content published online — that is considered to threaten public order, communal harmony, or national security. Critics of the vagueness of such provisions, their overbreadth, and susceptibility to misuse have observed that the negative effects may extend to the chilling of legitimate speech and dissent — especially when the latter is applied to political, religious or minority views.[7] (Kumari, P, 2025)

Nevertheless, the simultaneous enforcement of several legal instruments, namely the IT Act, the IT Rules, and *BNS* provisions, causes a multitude of problems in both legal and practical spheres. There is a considerable chance for duplicacy, gaps in application, and imposition of excessive regulations. The intermediaries might have to deal with the issue of conflicting responsibilities in relation to the IT Rules and criminal law. Removal under Section 69A (or notice under Section 79) might be done on the bases that are broad or vague while the criminal prosecution under *BNS* provisions will constitute an additional layer of the punitive threat. The system, so to speak, gives birth to over-censorship, self-censorship by platforms, and the creation of a hostile environment for the rightful expression. (Centre for Internet and Society, 2021)

[7] Kumari, P. (2025). Critical Analysis of Free Speech and Hate Speech on Digital Platforms. Advances in Consumer Research, 2(3).

**Table 1:**

*Key statutory provisions and their regulatory effects:*

| **Statutory Provision / Rule** | **Core Function / Effect** | **Potential Risks / Critiques** |
|---|---|---|
| **Section 69A, IT Act, 2000** | Government power to block access to online content for "sovereignty, security, public order, incitement to offence" . | Risk of executive overreach; lack of prior judicial oversight; possible arbitrary or politically motivated blocking. |
| **Section 79, IT Act, 2000** | Safe-harbour for intermediaries if they remove "unlawful content" upon notice. | Incentivises over-removal ("collateral censorship"); platforms may err on side of takedown to avoid liability. |
| **IT (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021** | Duties on intermediaries: grievance redressal, record-keeping, traceability, notice-and-takedown, compliance mechanisms. | Overbroad definitions; privacy concerns; procedural flaws; excessive delegation of content regulation; chilling effect on speech. |
| **BNS Sections (194, 297, 351, sedition, etc.)** | Criminal liability for hate speech, incitement, public mischief, religious insult, sedition, etc., as applied online. | Vague language, overbroad scope, potential misuse, targeting dissent or minority voices; deterrence of lawful criticism or debate. |

### 3.4 Overlap, duplication, and inter-statutory conflicts

One more thing besides duplication is the question of the constitutional legitimacy of the regulatory provisions concerned. For instance, the IT Rules, 2021 have been

termed as ultra vires the IT Act, as they bring in huge new terms such as "digital media" or "news and current affairs" which are not mentioned in the parent legislation; this raises serious doubts about the power of delegated legislation and due process.[8](Ashwini, S, 2021)

The move to platform-based, private moderation indicates that content governance is increasingly taking place outside the judicial realm. Intermediaries become the de facto gatekeepers who in effect decide what speech will survive — usually without any transparent criteria or judicial reviews. This situation, which sometimes goes by the name of "new-school regulation," depicts a fundamental change in the manner speech is controlled: not always through direct state intervention, but by regulatory powers over private firms that might over-comply to cut their liability. (Variath, B, 2023)

The statutory and regulatory framework for online speech in India — consisting of the IT Act, the 2021 IT Rules, and existing criminal law under the *BNS* — presents a complex, multifaceted interface of duties, boundaries, and liabilities that overlap and interact. Even though some regulation might be considered constitutionally justified to some extent under Article 19(2), the wide latitude, ambiguity, and procedural darkness of many provisions would still be a source of great risks for free speech, dissention, and democratic discussions.

The challenges presented indicate a strong case for reform in the doctrine that would draw the line clearly on what restrictions are allowed, provide more stringent language in statutes (especially the term "unlawful content"), and guarantee rights for the accused (i.e., the right to a notice, hearing, and appeal), etc. The only way through such reforms is the creation of a statutory framework that will be virtually aligned with the constitutional guarantee of free speech for the digital era.

## 4. HATE SPEECH AND CRIMINAL LAW IN THE ONLINE SPHERE

### 4.1 Definitional problems: what is "hate speech" under Indian law?

[8] Ashwini, S. (2021). Social media platform regulation in India–A specialreference to the information technology (Intermediary Guidelines and Digital Media Ethics Code) rules, 2021. Perspectives on Platform Regulation, 215-232

Hate speech is a term very hard to define in India as there is no independent statutory definition for it, rather, the control of the use of extreme or derogatory expressions aimed at certain groups is scattered among several criminal laws such as BNS Sections 194 (causing enmity among different groups), 195 (making statements harmful to national integration), 297 (committing acts that are deliberately and maliciously aimed at outraging religious feelings), 351 (making statements that could result in public disorder) and other public-order crimes — each of these categories having different penalties and proof levels. Both the Law Commission (No. 267) and the Supreme Court in Pravasi Bhalai Sangathan have pointed out that the lack of a single statutory definition not only creates doctrinal uncertainty and enforcement inconsistency but also leads to the situation where a common man's conception of "hate" might not necessarily fall within the boundaries of any specific BNS offence while, on the other hand, some BNS provisions are so wide that they could include even strong political or artistic expression when interpreted broadly. The division of authority among various jurisdictions and the existing infractions have led to the situation where it is necessary for the authorities to convert a moral idea of "hate incitement" into the specifics of the entrenched crimes – an ordeal even more difficult if the supposed speech is the one that has been posted online and its reach increased by the platform's algorithms. (Pravasi Bhalai Sangathan v. Union of India, 2014)

### 4.2 Elements, mens rea and statutory vagueness—risks of over-breadth and chilling effect

The *BNS* sections on substantive elements and mens rea differ significantly: Section *194* focuses on acts that arouse discord or hatred among communities (and is usually prosecuted without the requirement of proving a specific violent outcome), whereas Section *297* deals with intentional and malicious actions aimed at causing outrage among the adherents of a particular religion. The difference between "intent" or "knowledge" and lower standards, such as negligence or mere recklessness, is decisive; laws written in vague terms (e.g., "outrage religious feelings," "promote enmity") give the police major power to file FIRs even when there is no immediate threat of violence. In the context of the Internet, the rapid dissemination of posts and their wide sharing lead to complaints and FIRs

multiplying at a fast pace, resulting in a large number of prosecutions with low conviction rates—this pattern can be seen in NCRB data indicating a significant increase in offences registered under Section *194* between 2014 and 2020 (from 323 to 1,804 cases, an increase of about 458.51%). The rise in numbers, along with the low conviction rates reported in some studies, points to the situation where numerous cases are initiated but do not meet the proof requirement for conviction; nonetheless, the process itself suppresses speech.[9]

| **Year** | **Cases registered under *BNS* §*194*** | **% change from 2014** | **Source** |
|---|---|---|---|
| 2014 | 323 | — | NCRB via Business Standard. https://www.business-standard.com/article/current-affairs/500-rise-in-cases-filed-under-hate-speech-law-in-seven-years-ncrb-122062600627_1.html |
| 2020 | 1,804 | +458.51% | NCRB via Business Standard. https://www.business-standard.com/article/current-affairs/500-rise-in-cases-filed-under-hate-speech-law-in-seven-years-ncrb-122062600627_1.html |

**(Source: Statistics derived from NCRB reporting as summarised in Business Standard, 2022)**

A different dataset that serves as a second example illustrates the geographical location of §*194* filings in 2020, namely, Tamil Nadu (303, 16.8%), Uttar Pradesh (243, 13.47%), Telangana (151, 8.37%), Assam (147, 8.15%), and Andhra Pradesh (142, 7.87%). The concentration of these states reflects not only the local political dynamics but also the law enforcement strategies. Furthermore, it indicates that

[9] Business Standard. (2022, June 27). 500% rise in cases filed under hate-speech law in seven years — NCRB

prosecutions for online speech are not uniformly distributed and might, to some extent, be influenced by regional identity politics. The concentration of prosecutions in particular regions has raised questions about the equal application of law and that the use of criminal law is selective to silence dissent or minority voices.

| State | No. of §*194* cases (2020) | % of national total (1804) | Source |
|---|---|---|---|
| Tamil Nadu | 303 | 16.80% | Business Standard / NCRB summary. https://www.business-standard.com/article/current-affairs/500-rise-in-cases-filed-under-hate-speech-law-in-seven-years-ncrb-122062600627_1.html |
| Uttar Pradesh | 243 | 13.47% | Business Standard / NCRB summary. https://www.business-standard.com/article/current-affairs/500-rise-in-cases-filed-under-hate-speech-law-in-seven-years-ncrb-122062600627_1.html |
| Telangana | 151 | 8.37% | Business Standard / NCRB summary. https://www.business-standard.com/article/current-affairs/500-rise-in-cases-filed-under-hate-speech-law-in-seven-years-ncrb-122062600627_1.html |
| Assam | 147 | 8.15% | Business Standard / NCRB summary. https://www.business-standard.com/article/current-affairs/500-rise-in-cases-filed-under-hate-speech-law-in- |

| | | | |
|---|---|---|---|
| | | | seven-years-ncrb-122062600627_1.html |
| Andhra Pradesh | 142 | 7.87% | Business Standard / NCRB summary. https://www.business-standard.com/article/current-affairs/500-rise-in-cases-filed-under-hate-speech-law-in-seven-years-ncrb-122062600627_1.html |

**(Source: Figures and percentages computed from NCRB totals as reported.)**

### 4.3 Case law

In the case of *Romesh thappar v. State of Madras*, *1950*[10], the Supreme Court declared a ban order on an unconstitutional publication. The Court's reasoning was that the rights under § 19(1)(a) were violated since there were no justifications presented to enact the pre-publication restrictions. This case set a standard for every restriction on prior restraint and even the whole area of free speech.

Then, in *Bennett Coleman & Co. v. Union of India (1973)*[11], the Court discarded the newsprint policy, an instrument of quantitative controls over the press, saying that Article 19 has the only absolute and unqualified freedom of the press with its two dimensions: qualitative and quantitative; the ruling limits government measures that threaten the flow and diversity of opinions.

It was a case of *S. Rangarajan v. P. Jagjivan Ram, (19*89)[12] which wielded the power of the Court in a film-censorship and expression ruling that stemmed from the Court through licensing/censoring of cinematic content must apply proportionality and also context must be considered; the ruling secures political and artistic speech from excessive regulatory oppression under Article 19(1)(a).

Moreover, in *Kedar Nath Singh v. State Of Bihar,* (1962)[13], the Supreme Court

[10] *Romesh Thappar v. State of Madras*, A.I.R. 1950 S.C. 124
[11] *Bennett Coleman & Co. v. Union of India*, (1973) 2 S.C.C. 788
[12] *S. Rangarajan v. P. Jagjivan Ram*, (1989) 2 S.C.C. 574
[13] *Kedar Nath Singh v. State of Bihar*, (1962) Supp. 2 S.C.R. 769, A.I.R. 1962 S.C. 955

declared the sedition law (*BNS §150*) constitutional but at the same time limited its scope by confining only those acts that are intended or have a tendency to incite violence or public disorder thus setting a proximate-harm barrier to dissent criminalization.

Finally, in the case of *Pravasi Bhalai Sangathan v. Union Of India & ORS., (2014)*[14], the Court acknowledged the public damage caused by hate speech and suggested the Law Commission to look at the problem of the statutory definitions, and at the same time, emphasized that the existing penal provisions such as *BNS §§194*, *297*, *351* should be strictly enforced; however, the Court warned against the introduction of vague new offences that might suppress free speech.

*Shreya Singhal v. Union Of India, (2015)*[15] The apex court declared Section 66A of the IT Act, which imposed restrictions on free speech, null and void, relying on the argument of vagueness and chilling effect. Then again, it accepted Section 69A and the procedural blocking framework as well, thereby determining the limits of free speech online and of the duty of intermediaries (Section 79).

*Anuradha Bhasin v. Union Of India*, *(2020)*[16] The Court, while dealing with the issue of net shutdowns in Jammu & Kashmir, stated that the government order regarding blanket suspension implicates the right to free speech under Article 19(1)(a) and, therefore, should have the backing of legal safeguards and be subject to judicial review; the ruling has curtailed the power of the government to disrupt crucial communication that has an impact on press and public speaking.

*Anvar P.V. v. P.K. Basheer & ors., (2014)* [17]The apex court gave guidelines regarding the use of electronic documents as evidence in court under Evidence Act Section 65B by specifying the requirements for proving computer outputs; the ruling has a significant impact on the prosecution and defense in online speech and cyber crime cases where digital authenticity is questioned.

[14] Pravasi Bhalai Sangathan v. Union of India, (2014) 11 S.C.C. 477
[15] Shreya Singhal v. Union of India, (2015) 5 S.C.C. 1, A.I.R. 2015 S.C. 1523
[16] Anuradha Bhasin v. Union of India, (2020) 3 S.C.C. 637
[17] Anvar P.V. v. P.K. Basheer, (2014) 10 S.C.C. 473

*Avnish bajaj v. State (NCT of Delhi), (2005) (BAZEE.COM CASE)*[18] The situation was indeed very complicated and thus the Delhi High Court was forced to consider whether to place the liability on the intermediates (for their complicity) or on the users (for the upload) when the users uploaded obscenities under the IT Act sections 67/85 and the I.P.C. Provisions. This incident sparked a discussion on the gatekeeper liability and moreover, it affected the safe harbor doctrines and regulations for intermediaries later (Section 79). **J***ustice K.S. Puttaswamy (RETD.) v. Union Of India, (2017)* [19]The nine-judge Court granted the right to privacy under Article 21, a now very basic principle in the adjudication of data-traceability, data-access, and Surveillance demands (e.g. compulsory disclosure under tracing rules in the IT (Intermediary) Regulations), thus privacy limits delineate any prohibition on online speech.

### 4.4 The regulation of hate speech

In particular, should consider political speech and dissent as protected categories for the sake of democracy since the democratic process needs robust critique. Therefore, the constitutional exception doctrine must require (i) a clear and legitimate aim, e.g., preventing imminent violence or protecting vulnerable groups, (ii) close connection between the speech and the harm (not mere offence), and (iii) a specially crafted mens rea (intent or knowledge rather than negligence). Overbroad laws that equate "hurt feelings" or "offence" with crime without the requirement of proximate harm risk silencing political discourse. The online sphere exacerbates the situation: expressions that use hyperbolic language, parody, or academic critique can be considered criminally actionable even when reused in a different context. In this way, every exception must be carefully defined in terms of its limits, require the state to prove the existence of a close connection between the harm and the speech, and uphold the procedural protections (such as notice, the opportunity to be heard, and speedy adjudication), while reserving the criminal law for the most serious, violence-provoking speech only.[20]

[18] *Avnish Bajaj v. State (NCT of Delhi)*, 116 (2005) D.L.T. 427
[19] *Justice K.S. Puttaswamy (Retd.) v. Union of India*, (2017) 10 S.C.C. 1
[20] Centre for Internet and Society. (2021). On the legality and constitutionality of the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021

## INTERMEDIARY LIABILITY, PLATFORM GOVERNANCE AND PRIVATE MODERATION

### 4.5 Safe-harbour regimes and conditional immunity (procedural preconditions)

The IT Act's Section 79 offers intermediaries a partial safe harbor protection for third-party content, with a catch: they must perform "due diligence" and take action when they gain actual knowledge or when there are court/government orders. The IT Rules, 2021 set out these duties by assigning the role of grievance officers, imposing time-limited takedown, and maintaining records of "significant social media intermediaries" among other things. This conditional immunity mechanism puts a lot of regulatory responsibility on the intermediaries: compliance gives immunity, non-compliance results in liability. Though this encourages platforms to adopt notice-and-takedown systems, it might also lead to "over-cautious" behavior (for example, platforms removing content that is only slightly objectionable in a preemptive manner) because the risk of being non-compliant is civil/criminal exposure. (MeitY, 2021 Rules)

### 4.6 Notice-and-takedown, automated moderation, and transparency obligations

Notice-and-takedown is still the main operational model: the intermediaries are obliged to remove content within the specified time once a complaint or a government order comes their way. Automated moderation (like filters, hash-matching, machine-learning classifiers) is very popular because it can be done quickly and at high volumes, but such automation cannot really understand the context and hence, a lot of false positives are generated. The IT Rules demand transparency through (compliance reports at regular intervals, grievance redress) but they do not provide for independent supervision and effective appeal channels; without court or legislative review, removal judgments may be concealed and in effect, irrevocable.

### 4.7 Decentralization, encryption, and platform incentives issues.

Technical design choices — end-to-end encryption, ephemeral messaging, and decentralized networks — make it difficult for notice, traceability, and law

enforcement access. Also, the requirement to trace the originators (traceability) and the users' privacy rights conflict with each other as they are the opposite sides of the Article 21 values. The commercial motivations of the platforms (the user's engagement) are also in support of the algorithmic promotion of the lots of polarizing content; hence, the regulatory obligations of the intermediaries must be complemented by the measures to protect against the mission creep (surveillance, erosion of encryption) and also to decide reasonably on how much access to data is necessary. (Puttaswamy, 2017 — privacy anchor)

### 4.8 Remedies, appeals and independent oversight mechanisms

An intermediary regime that respects rights needs a statutory notice, a time-limited counter-notice procedure, independent appellate review (tribunal or judicial fast-track), transparency reports and civil-society monitors' audit rights. In the absence of such mechanisms, content moderation transfers public law powers to private entities without proper checks, leading to a structural problem for constitutional accountability. Recent policy proposals and judicial directions (for instance, those calling for registration and enforcement in hate-speech orders) highlight the urgency of procedural safeguards being embedded alongside platform obligations. (PRS Legislative Research, 2021)

## 5. CONSTITUTIONAL TESTS & ANALYTICAL FRAMEWORKS FOR ONLINE REGULATION

### 5.1 Reasonable restriction test: legitimate aim, proximate nexus, and proportionality

Any limitation imposed on Article 19(1)(a) must comply with the conditions set out in Article 19(2) the restriction must be directed towards a legitimate purpose (for instance, public order, security), have a direct link to that purpose, and be proportionate. In the case of online activities, the application of proportionality involves easing the restrictions: the least-restrictive methods, time and area limitations where possible, and undeniable procedural safeguards. Governments should thus delineate the specific wrongs they aim to combat (such as incitement to imminent violence, organized calls for ethnic cleansing, etc.) instead of using

broad categories like "offensive" speech. (Constitution of India, 1950)

### 5.2 The harm principle, strict scrutiny vs. reasonableness in the Indian context

Indian jurisprudence has preferred a reasonableness standard instead of an absolutist First Amendment strict scrutiny; nonetheless, the harm principle (speech may be restricted if significant, demonstrable harms are to be prevented) has been and still is a key aspect. For online content regulation this means that deleting content or charging people with a crime must be associated with evidence of likely real-word harm or serious risk to public order; without such an association, the restrictions will not pass the tests of both proximate nexus and proportionality. The legal principle should further guard against imposition of stringent procedural protections in instances wherein the claimed harm has to do with political speech, satire, or academic debate. (CJP Team, 2025)

### 5.3 Vagueness and chilling effect analysis

The vagueness doctrine closes the door to statutes that either do not provide citizens with a notice that is fair regarding the conduct that is prohibited or that allow arbitrary enforcement. Courts will not tolerate online regulation that employs non-specific terms (like “unlawful information”, “harmful content”... without definitional clarity) and hence will either strike such regulations down or limit them. Such vagueness in the law leads to users and platforms preemptive censoring as they cannot accurately foresee legal risk. The doctrinal cure is effective textual precision (with defined elements, mens rea, procedural remedies) and institutional safeguards (independent review, time-limits, and transparency) which will lower the risks of both overbreadth and arbitrary enforcement.[21]

## 6. COMPARATIVE PERSPECTIVES

### 6.1 United States: Section 230 and First Amendment contrasts

The United States model is based on two pillars that complement each other: a strong First Amendment tradition and the statutory liability immunity of

[21] Legitimate India. (2025). Balancing freedom of expression and hate speech: A critical analysis of India’s IT Act, 2000

intermediaries, which is 47 U.S.C. § 230. Section 230 provides a very broad liability shield for the most diverse content uploaded by users as well as for platforms that are acting "in good faith" moderation-wise; thus, it positions both a high level of expressive freedom and private moderation without affecting the role of platforms as de facto publishers. U.S. constitutional law—demonstrated in cases such as *Reno v. ACLU (1997)*[22] and *Packingham v. North Carolina,(2017)*[23]—on the other hand, not only breathes a new life into the internet as the most liberated area of expression but also imposes such strict standards on Government that it has to resort to very precise and narrowly focused measures each time it intends to stifle the very expression it seeks to prohibit. The recent Supreme Court ruling in *Gonzalez v. Google LLC, (2023)*[24] examined the extent of a platform's liability for suggestion made through algorithms but at the same time preserved the major part of §230 immunity, and so it gave a sign that it was a bit of a judge's disagreement around the issue of transferring huge content moderation from Congress to the courts. The U.S. method is different from that of India, which has given a huge burden of compliance coupled with a conditional safe-harbor model under Section 79 of the IT Act, 2000, where intermediary immunity is simply condition upon adherence to due-diligence obligations and statutory practices. In addition, governmental powers of takedown (e.g., Section 69A), and rules (IT Rules, 2021) put platforms in a more direct regulatory control through the government. (Cornell Law, 2024)

### 6.2 European Union: Platform regulation, hateful content regulation and human-rights balancing

The regulatory system of the European Union, which reached its peak in the form of the Digital Services Act (DSA) and connected instruments, is built on the prescriptive, duty-based approach. Very large online platforms must perform specific functions including the removal of illegal content, risk-assessment, transparent reporting, and algorithmic supervision, while the enforcement of national laws against hate speech and incitement remains a function of the Member States. The DSA tests a human-rights-based—the platform duties must be

[22] Reno v. ACLU, 521 U.S. 844 (1997)
[23] Packingham v. North Carolina, 582 U.S. (2017)
[24] Gonzalez v. Google LLC, 598 U.S. (2023)

consistent with the essential rights that include the right to free speech and the right to privacy and data protection—and it imposes novel responsibilities for the diffusion of systemic risks and independent evaluation of the recommender systems. Meanwhile, the seasoned EU efforts (e.g., the Code of Conduct on countering illegal hate speech online and the Audiovisual Media Services Directive) ensure that clearly illegal hate speech is removed quickly while providing lawful expression with procedural safeguards. In contrast to the U.S. permissive immunity model and India's conditional-compliance model, the EU's approach emphasizes mandatory platform obligations supported by administrative oversight and cross-border enforcement. (European Commission, 2024)

### 6.3 Lessons for India: tailored transplant vs. constitutional fit

From the comparative law perspective, there is no single "best" solution and India has to adjust the regulatory measure according to the constitutional values. The U.S. model enables a very robust protection for speech and shows utmost respect to private moderation, but its hands-off approach to online platforms is constitutionally based on the First Amendment and the legislative history of §230—an exact replication would be both unconstitutional and politically difficult in India where Article 19(2) allows for restrictions and where the government already claims stronger regulatory powers. The EU's duty-oriented approach is supportive of the technical tools—systemic risk assessments, transparency requirements, and procedural safeguards for content removal—which coincide with India's demand for accountability but need to be reshaped according to the constitutional provisions in India (Article 19(1)(a), Article 19(2), Article 21) so that they do not lead to excessive restriction. In practical terms, the country of India is opened to the following: (a) keeping conditional intermediary immunity (Section 79) but improving due process guarantees (time-limited appeals, independent oversight); (b) instituting selective transparency and audit requirements (modelled after the DSA) while providing for privacy and encryption that are consistent with Puttaswamy (2017); and (c) restricting the scope of criminal law (BNS Sections 194, 297, 351) through stipulating clearer mens rea conditions so that criminal law touches only the proximate, violent or seriously harmful acts and not mere annoyance. (MeitY; Puttaswamy, 2017)

## 7. EMERGING TECHNOLOGIES AND NEW CHALLENGES

### 7.1 Deepfakes, synthetic media, and attribution problems

The popularity of deepfakes and synthetic media is shaking up the old evidentiary rules and making it harder to assign responsibility for the content: fake video and sound can now be made at a very low cost and distributed over a large area, thereby exposing people to real risks of loss of reputation, misleading information being circulated, and disturbances in the voting process. The current laws do not specifically deal with the issue of synthetic media harms; though the IT Act's Section 69A and some criminal provisions (like *BNS* sections 499/500 on defamation, and *351* on public mischief) could be called upon, those provisions were not written with the idea of machine-generated lies. The existing situation is such that there is a need for regulatory changes to supplement the existing law—definitions of "synthetic media," specific mens rea for different types of synthetic media, and quick forensic procedures—together with measures to protect against the chilling effect on legitimate parody, satire, or political speeches. Moreover, technology creates new challenges for attribution: the requirement of tracing the originator set in the IT Rules, 2021 clash with encryption and pose privacy issues under Article 21 unless the disclosure is narrow and controlled by court standards and judicial supervision. (MeitY, 2021)

### 7.2 AI moderation, automated removal, and contestability of algorithmic decisions

To manage the enormous volume of content on a scalable level, the use of automated moderation (hash-matching, classification, recommender algorithms) is a must, but this also leads to many instances of false accusations and "black box" decision-making. The question of the constitutionality of algorithmic removals from both procedural and substantive perspectives arises: when the removal is mandated by law, is it still “state action and what are the due-process mechanisms ensuring that the removal can be contested? The EU’s DSA imposes a right of appeal and transparency obligations on algorithmic systems while India’s IT Rules require grievance officers and transparency reports but do not yet impose independent auditability of models or clear appeals structures that bind the

Platform and the State. The AI-driven moderation can result in de facto prior restraint and unequal speech treatment without strong counter-notice procedures, time-limits, and audit rights. Regulatory measures should accordingly provide for explanation-ready logs, human review thresholds for grey content, and statutory rights to speedy administrative or judicial review. (DSA / MeitY, 2021)

### 7.3 Cross-border enforcement and Jurisdictional Limits

The cross-border nature of the Internet makes Mosaic enforcement: immediate local takedown of content hosted on foreign servers might not be possible and might require mutual legal assistance, transnational notice mechanisms, or law-to-platform channels (e.g., trusted flaggers) under provider cooperation. While the EU's DSA provides for harmonized procedures across Member States, India has to rely on platform global policies plus Section 69A blocking orders. However, unilateral takedowns risk the exporting of local censorship norms while weak cross-border cooperation allows the prolongation of harmful content. Therefore, India should go for a cooperative strategy with others that will expedite the elimination of violent/terrorist propaganda and child-sexual-abuse material while opposing the broad cross-border application of vague domestic offenses that would otherwise undermine global expression norms.

## 8. IMPLEMENTATION CHALLENGES AND ENFORCEMENT REALITIES

### 8.1 Police Practice, Prosecutorial Discretion, and Misuse Risks

Implementation has a lot to do with the policing of practice just like it does with statutory design. The police and the prosecutors on a regular basis make use of the *BNS* provisions (Sections *194*, *297*, *351*) and the complaint mechanisms of the IT Act to file FIRs in online speech cases; the use of vague statutory language has led to the granting of wide prosecutorial discretion and the selective enforcement of the law, which has in turn led to the disproportionate criminalization of dissidents, journalists, and minority voices. The courts have on numerous occasions reiterated the need for safeguards, but the first-line officers still detain or file cases without carrying out proper threshold assessments, thereby producing a chilling effect. Thus, the reforms of the institutions have to be directed towards the establishment

of guidelines for the prosecution, the requirement of prima facie review before arrest in speech matters, and the provision of training for cyber forensic and proportionality assessment. (Shreya Singhal, 2015; Puttaswamy, 2017)

### 8.2 Capacity limitations, resource concerns, and evidential difficulties online

The digital forensics, log-preservation, the chain-of-custody, and specialist prosecutors are the main requirements of the online harm investigations; the inadequacy of police units is one of the common reasons for such units taking long and leading poorly substantiated prosecutions. Evidence fragmentation (deleting posts, ephemeral messaging) and hosting in different countries make it more difficult to proving the origin and the intention. The record-keeping and preservation requirements in the IT Rules (for intermediaries) are helpful but they also impose a considerable burden on both the platforms and investigators in terms of resources; thus, fiscal and institutional investments in cyber-crime units that are specialized and have quick mutual-legal-assistance channels are essential. In addition, the enforcement of procedural safeguards (judicial authorization for invasive data access, time-limited warrants) will be necessary to avoid mission-creep and privacy erosion. (MeitY, 2021)

### 8.3 Remedies for wrongful takedown and strategic litigation against public participation

Takedown of content by mistake and strategic lawsuits against public participation (SLAPPs) cause serious damages that are closely related to the enforcement of the law. The platforms' responses to government or user complaints concerning takedowns usually do not provide any good counter-notice pathways; the people who have been wrongly removed suffer loss of reputation and income, and their access to justice is very limited. Under India's IT Rules, grievance officers and appellate grievance committees are a must, but on the other hand, there are no independent and quick adjudication mechanisms (such as statutory tribunals or fast-track courts) in place that would allow to award damages, restore the content and punish the complainants who are causing trouble. Data from research indicate that there is a large number of user complaints in India, for instance, there were 76,321 complaints by Indian users in June 2023 across different intermediaries

(according to platform transparency reports), whereas 55,497 national government requests to Meta were made in the first half of 2022 with about 66.6% compliance — these figures demonstrate the extent of moderation and government interference. The trends point out the importance of statutory counter-notice procedures, anti-SLAPP protections, and remedial damages to prevent the misuse of rights. (Medianama, 2023; Times of India/Meta data, 2022)

| **Platform / Metric** | **Number (period)** | **Percentage / Note** | **Source** |
|---|---|---|---|
| Indian user grievances (June 2023) | 76,321 | — | MediaNama reporting of platform transparency reports. https://www.medianama.com/2023/08/223-india-social-media-user-grievances-june-2/ (MediaNama, 2023). |
| Government requests to Meta (India, H1 2022) | 55,497 requests | 66.59% compliance (data provided) | Times of India summary of Meta Transparency Report. https://timesofindia.indiatimes.com/gadgets-news/meta-transparency-report-government-requests-for-user-data-increases-in-india-globally/articleshow/95739562.cms (Times of India / Meta, 2022). |

The problems surrounding the use of comparative analysis, legal innovation brought in by new technology, and different levels of implementation failures in terms of maintaining constitutional rights, have led the Indian regulatory

framework to rely on the following: Doctrinal clarity (well-defined offenses and mens rea), institutional architecture (independent appeal bodies, forensic capacities), procedural protections (counter-notice, judicial authorization), and the technical safeguard of the privacy-protecting traceability frameworks. The prescribed statutory tools (IT Act Sections 69A, 79; IT Rules, 2021; *BNS* provisions) are aligned with the reasonableness tests of Article 19(1)(a) and Article 19(2) so that online regulation does not only suppress the most serious violations but also protects the democratic discourse. (MeitY, 2021; Shreya Singhal, 2015)

## 9. RECOMMENDATIONS

While dealing with the negative impacts of the internet, to ensure that people's freedom of speech is protected in the digital world, the government should (a) make the laws clearer and more precise: change the Information Technology Act 2000 to give a definition of "unlawful content" and set the mens rea thresholds (intent/knowledge) for the commission of offenses punishable by criminal sanctions, and also change the *BNS* Sections that deal with incitement to religious hatred (Section *194*), insult to religions (Section *297*), and spreading false rumors likely to cause disturbance (Section *351*) to consider harm and intent rather than the vague criteria of “outrage” or “promotion”; (b) keep the conditional intermediary immunity under Section 79 but re-mold the procedural duties in the IT (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 into a statutory due-process regime that requires notice, limited-time takedown, counter-notice, and an independent appeal to a fast-track tribunal. (c) Legislate the very limited, privacy-aware tracing powers only with judicial approval in order to make the blocking powers of Section 69A and the requirements of traceability compatible with the privacy guarantees of Article 21 established in the case of Justice K.S. Puttaswamy (2017).

It is important to develop institutional capacity and transparency: establish a digital-forensics and public-order assessment cell in state police with compulsory prima-facie review procedures prior to the registration of FIRs in matters of online speech; impose upon platforms the obligation to release transparency reports and independent audit results (including accuracy/false-positive rates of automated moderation) in a machine-readable format; and incorporate anti-SLAPP protections and statutory remedies (restoration, damages, and penalties for vexatious complainants) to deter misuse of

takedown and criminal complaints.

It is advisable to introduce narrowly defined, technology-specific offences (e.g., the intentional distribution of synthetic media for the purpose of inciting violence) and multi-lateral cooperation mechanisms for the removal of content across borders based on EU/Digital Services Act best practices but adapted to India's constitutional requirements under Article 19(2).

## 10. CONCLUSION

The right to free speech guaranteed by India's Constitution through Article 19(1)(a) must be upheld regardless of the State's and society's encounter with digital-related threats. The interpretative guidance of Article 19(2) — restricting "sensible limitations" only for stated purposes including public order and state's safety — demands that any legislative or regulatory measures be very specific, proportionate, and have the strongest possible procedural safeguards. In reality, though, the legal framework formed by the Information Technology Act, 2000 (e.g., Sections 69A and 79), the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, and the overlapping criminal laws in the *Bharatiya Nyaya Sanhita* (Sections *194*, *297*, *351*, and others) creates a complex and confusing system of potential regulation that could suppress free speech, result in the private platforms or the police dealing with matters without proper safeguards, and the courts being pushed to the sidelines. (Shreya Singhal, 2015)

From the judicial interpretation, the courts take on the role of the essential guards. The Supreme Court's ruling in the case of Shreya Singhal v. Union of India, (2015), which nullified Section 66A and stressed the chilling impact of imprecise penal rules, confirmed that limitations on the expression over the internet should be very high in terms of certainty and harm proximity. Likewise, the Court in the case of Pravasi Bhalai Sangathan v. Union of India, (2014) identified the social negative effects of hate speech and asked for a re-evaluation of the definitions and remedies by the legislature. The nine-judge verdict in Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) firmly places Privacy (Article 21) as a constitutional value which should govern the rules of traceability and data access. All these authorities combined together speak of a constitutional synthesis: the principles of free expression, privacy and rule-of-law must be the guiding

factors in any regulation of online-speech. (Pravasi Bhalai Sangathan, 2014)

The operational policy choices result from this synthesis. To begin with, it is important to have statutory precision: create clear definitions in the statutes (for "hate speech", "unlawful information", and "synthetic media"), require mens rea of either intent or knowledge whenever criminal punishment is imposed, and introduce proximate-harm gates to protect the criminal law from being applied just for offence or hurt feeling. The present notice-and-takedown mechanism of the IT Rules should be transformed into a statutory, time-bound due-process frame so that the immunity under Section 79 is based on adherence to clear notice, counter-notice and independent appeal routes rather than the resolution left to non-transparent platform policies. The State should conduct prior internal review when using Section 69A blocking powers, and it should be followed by prompt judicial oversight to prevent over-enforcement by the executive. (Puttaswamy, 2017).

Two essential components are institutional and procedural safeguards to stop misuse. Police and prosecutors must set prima-facie thresholds before FIR registration in speech-related cases and undergo training on digital evidence and proportionality assessment; otherwise, they will have an increasing number of low-merit complaints, which will create chilling effects that will not go away even if the defendant is eventually cleared. The platforms would have to make their machine-readable transparency data available, implement fast counter-notice procedures, and permit independent audit of AI moderation systems to ensure that they have safe-harbour protection. Anti-SLAPP laws and statutory redress (restoration orders, compensation, and fines for dishonest complainants) will discourage litigation for the purposes of rights' infringement and unwarranted retakes that silence the voice of the public. (Pravasi Bhalai Sangathan, 2014)

The third point is that the law should respond to technology with targeted measures instead of wholesale prohibitions. Offences should be created only for the most dangerous technologies when used with malice (e.g., fake videos made for the purpose of inciting violence or large-scale fraud) while still allowing for satire, parody, academic critique, and political discourse. The traceability requirements should be limited, judicially authorized, and data-minimizing so that they do not interfere with privacy rights under Article 21 as stated in Puttaswamy. Enforcement across borders should

depend on notice systems being able to work together and on international collaboration for violent extremist content and child sexual abuse material, but India should not allow its ambiguous domestic principles to affect global freedom of expression. (Shreya Singhal, 2015)

To sum up, the constitutional command is unambiguous: secure the essence of Article 19(1)(a) and allow only those restrictions which are proven to be necessary, proportionate, and following the due process of Article 19(2) at the same time. Law reform must then aim at clarity, procedural due process, independent oversight, technical specificity, and institutional capacity-building. A new regime that limits the application of criminal law (BNS Sections 194, 297, 351), clarifies the obligations of intermediaries under Section 79, ties up the government's blocking under Section 69A with judicial review, and acknowledges privacy (Puttaswamy) will be able to synchronize the democratic discourse with the genuine needs of public order and safety during India’s digital age.

## REFERENCES

1. Constitution of India. (1950). *The Constitution of India*. https://www.constitutionofindia.net/articles/article-19-protection-of-certain-rights-regarding-freedom-of-speech-etc/

2. Ministry of Electronics & Information Technology. (2021). *The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021* (updated 06.04.2023). Government of India. https://www.meity.gov.in/static/uploads/2024/02/Information-Technology-Intermediary-Guidelines-and-Digital-Media-Ethics-Code-Rules-2021-updated-06.04.2023-.pdf

3. Ministry of Law and Justice. (2000). *The Information Technology Act, 2000*. Government of India. https://www.indiacode.nic.in/bitstream/123456789/13116/1/it_act_2000_updated.pdf

4. Government of India. (1860). *The Bharatiya Nyaya Sanhita, 1860* (selected sections: *194*, *297*, *351*). https://www.indiacode.nic.in/repealedfileopen?rfilename=A1860-45.pdf

5. PRS Legislative Research. (2021). *The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 — PRS analysis*. https://prsindia.org/billtrack/the-information-technology-intermediary-guidelines-and-digital-media-ethics-code-rules-2021

6. The Economic Times. (2025). Centre moves to regulate deepfakes, AI media; MeitY proposes amendments to IT rules. https://m.economictimes.com/tech/artificial-intelligence/centre-moves-to-regulate-deepfakes-ai-media-meity-proposes-amendments-to-it-rules/articleshow/124737121.cms

7. Singh, D. P. (2025). An Analysis of the Article 19 (1)(a) and Article 19 (2) of the Indian Constitution and Distorting Form of Freedom of Speech and Expression in the Era of Social Media in India. *Available at SSRN 5100601*. https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=5100601

8. Centre for Internet and Society. (2021). *On the legality and constitutionality of the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021*. https://cis-india.org/internet-governance/legality-constitutionality-il-rules-digital-media-2021 (Centre for Internet & Society)

9. Srivastava, S. (2023). Freedom of Speech and Expression Scope of Article 19 (1)(a) in the Constitutional Framework and Reasonable Restrictions. *Issue 6 Int'l JL Mgmt. & Human.*, *6*, 1071. https://heinonline.org/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/ijlmhs26&section=96

10. Kumari, P. (2025). Critical Analysis of Free Speech and Hate Speech on Digital Platforms. *Advances in Consumer Research*, *2*(3). https://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=00989258&AN=185938520&h=bNKsMuj%2FRfXP%2FkeWu%2Fl05EX6Pw5fc2dv%2BFSojxOAZcNEZHXfBhVMGcksftP0OAh9az7cmgYwHqFYEugeSL%2Fx7Q%3D%3D&crl=c

11. Variath, B. (2023). Reviewing the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021. *Issue 2 Indian JL & Legal Rsch.*, *5*, 1. https://heinonline.org/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/injlolw11&section=517

12. Ashwini, S. (2021). Social media platform regulation in India–A special reference to the information technology (Intermediary Guidelines and Digital Media Ethics Code) rules, 2021. *Perspectives on Platform Regulation*, 215-232. https://library.oapen.org/bitstream/handle/20.500.12657/58180/external_content.pdf?sequence=1#page=215

13. NMIMS Law Review. (2021). *Part I: Communications Decency Act, Section 230 vis-à-vis Indian law on* Wadhwa, S. (2025). Content blocking orders and status of digital rights: Assessment of two key verdicts in India. *Information & Communications Technology Law*, *34*(1), 44-61*intermediary liability — A comparative analysis*. https://lawreview.nmims.edu/2021/04/14/part-1-communications-decency-act-section-230-vis-a-vis-indian-law-on-intermediary-liability/ (lawreview.nmims.edu)

14. Wadhwa, S. (2025). Content blocking orders and status of digital rights: Assessment of two key verdicts in India. *Information & Communications Technology Law*, *34*(1), 44-61. https://www.tandfonline.com/doi/abs/10.1080/13600834.2024.2406678

15. Amish Devgan v. Union of India & Ors., Writ Petition (Criminal) No. 160 of 2020; AIR (Online) 2020 SC 930. https://indiankanoon.org/doc/179868451/

16. Business Standard. (2022, June 27). 500% rise in cases filed under hate-speech law in seven years — NCRB. https://www.business-standard.com/article/current-affairs/500-rise-in-cases-filed-under-hate-speech-law-in-seven-years-ncrb-122062600627_1.html

17. Centre for Internet and Society. (2021). *On the legality and constitutionality of the* Information *Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021*. https://cis-india.org/internet-governance/legality-constitutionality-il-rules-digital-media-2021

18. Constitution of India. (1950). *The Constitution of India* (Article 19(1)(a), Article 19(2)). https://www.constitutionofindia.net/articles/article-19-protection-of-certain-rights-regarding-freedom-of-speech-etc/

19. NCRB data summary. (2022). Data dive: Sixfold rise in cases filed under hate-speech related law in 7 years. https://www.factchecker.in/data-dive/data-dive-sixfold-rise-in-cases-filed-under-hate-speech-related-law-in-7-years-822966

20. Law Commission of India. (2017). *Report No. 267: Hate Speech* (and related materials). (See summaries and references at: https://indianlawwatch.com/issue-of-hate-speech-law-commission-report-no-267/)

21. Pravasi Bhalai Sangathan v. Union of India & Ors., Writ Petition (Civil) No. 157 of 2013, (2014) 11 SCC 477; AIR 2014 SC 1591. https://opil.ouplaw.com/abstract/10.1093/law-ildc/2827in14.case.1/law-ildc-2827in14

22. PRS Legislative Research. (2021). *The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 — analysis*. https://prsindia.org/billtrack/the-information-technology-intermediary-guidelines-and-digital-media-ethics-code-rules-2021

23. Privacylibrary / Puttaswamy v. Union of India (2017). *Justice K.S. Puttaswamy (Retd.) vs Union of India*, (2017) 10 SCC 1. https://privacylibrary.ccgnlud.org/case/justice-ks-puttaswamy-ors-vs-union-of-india-ors

24. Shreya Singhal v. Union of India, Writ Petition (Criminal) No.167 of 2012; (2015) AIR SC 1523. Full text: https://indiankanoon.org/doc/110813550/

25. Shaheen Abdullah v. Union of India & Ors., Writ Petition (2022). Court documents and orders: https://www.livelaw.in/pdf_upload/872-shaheen-abdullah-v-union-of-india-21-oct-2022-440831.pdf

26. The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 (Ministry of Electronics & Information Technology). (2021; updated 06.04.2023). Government of India. https://www.meity.gov.in/static/uploads/2024/02/Information-Technology-Intermediary-Guidelines-and-Digital-Media-Ethics-Code-Rules-2021-updated-06.04.2023-.pdf

27. Legitimate India. (2025). *Balancing freedom of expression and hate speech: A critical analysis of India's IT Act, 2000*. https://legitimateindia.com/balancing-freedom-of-expression-and-hate-speech-a-critical-analysis-of-indias-it-act-2000/

28. Constitution of India. (1950). *The Constitution of India* (Art. 19(1)(a), Art. 19(2)).https://www.constitutionofindia.net/articles/article-19-protection-of-certain-rights-regarding-freedom-of-speech-etc/

29. Information Technology Act, 2000. (2000). Government of India. https://www.indiacode.nic.in/bitstream/123456789/13116/1/it_act_2000_updated.pdf

30. Ministry of Electronics & Information Technology. (2021). *The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021* (updated 06.04.2023). Government of India. https://www.meity.gov.in/static/uploads/2024/02/Information-Technology-Intermediary-Guidelines-and-Digital-Media-Ethics-Code-Rules-2021-updated-06.04.2023-.pdf

31. Shreya Singhal v. Union of India, (2015) 5 SCC 1. https://indiankanoon.org/doc/110813550/

32. Pravasi Bhalai Sangathan v. Union of India & Ors., (2014) 11 SCC 477. https://opil.ouplaw.com/abstract/10.1093/law-ildc/2827in14.case.1/law-ildc-2827in14

33. Justice K.S. Puttaswamy (Retd.) & Anr. v. Union of India & Ors., (2017) 10 SCC 1. https://cdnbbsr.s3waas.gov.in/s3ec0490f1f4972d133619a60c30f3559e/documents/aor_notice_circular/43.pdf

34. *Bharatiya Nyaya Sanhita*, 1860 (selected provisions: §§*194*, *297*, *351*). https://www.indiacode.nic.in/repealedfileopen?rfilename=A1860-45.pdf

35. PRS Legislative Research. (2021). *The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 — PRS analysis*.https://prsindia.org/billtrack/the-information-technology-intermediary-guidelines-and-digital-media-ethics-code-rules-2021.