IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Safe Harbor Provisions and Content Moderation: A Study of Legal Framework for Digital Media in India

Dr. Raghuvinder Singh, Professor (Retd.), Faculty of Law, Himachal Pradesh University, Shimla-05,

Mr. Sudhir, Research Scholar, Faculty of Law, Himachal Pradesh University, Shimla-05

Abstract

The sudden expansion of digital media in India has transformed communication, information exchange and public debate. Social media sites, online forums, and online news websites have become the hub of opinion making, debates and storytelling. India has one of the biggest social media populations in the world.[1] The digital media is a powerful tool for information exchange and unification of people across different audiences. But this digital expansion has also caused problems such as the spreading of misinformation, hate speech, and illicit content. To deal with such issues, the Indian legal system has put into place certain mechanisms like managing content and specifying the role of digital media intermediaries. This paper explores the concept of 'Safe-harbour' and role of intermediaries in content moderation. The paper highlights international framework, national framework, and judicial scrutiny over these issues of digital media and the challenges it poses to the present globalised world.

Keywords: Safe Harbour, content moderation, digital media

Introduction

The 'Center for the Study of Organized Hate' listed hundreds of instances of hate speeches by political parties in India during of 2024 elections. Of 1,165 hate speech incidents reported, 995 were first reported on social

[1] SIMON KEMP, DIGITAL 2025: INDIA, DataReportal 25 FEBRUARY 2025 available at: https://datareportal.com/reports/digital-2025-india (last visited on May 5, 2025).

media.[2] Facebook accounted for the 495 cases and YouTube accounted for 211 cases. Facebook took down only 3 videos, leaving 98.4% reported content available to the public domain.[3]The term 'safe-harbor' is often employed metaphorically to describe the way governments and legal authorities employ statutory instruments to hold local employees of international companies accountable for enforcing compliance with regulatory demands.[4] Ideally, accountability for social networking site content should rest with the authors who create it. Intermediaries are usually 'safe-harbor' insulated, sheltering founders and employees from criminal or individual liability, unless they actively engage or have special knowledge of the distribution of harmful content.[5] But more and more employees are threatened with personal liability for content moderation decisions. This has an adverse effect on free speech, as enshrined in the Constitution of India[6], because platforms can pre-emotively take down legitimate content in an attempt to avoid getting into legal difficulties. Over-policing content or zealous compliance with governmental directions are likely to destroy users' freedom of speech and gives rise to a culture in which fear of backlash trumps honest debate.

In the past few years, some powerful social media giants have withdrawn from some countries due to demands by governments for content moderation which go against free speech, rights to privacy, and global standards. Google's withdrawal from China[7], Telegram's bans in various countries[8], and X's (formerly Twitter) voluntary withdrawal from Brazil are a few examples.[9] All these are symptoms of the conflict between government control and the freedom of expression.

The Challenges of AI Content and Disinformation

The spread of disinformation and deep fakes have also caused governments to press platforms to assume responsibility.[10] The problem is most sensitive during crisis situations and election seasons, when AI-generated content presented as deepfakes can spread misinformation quickly. New legislations aimed at

[2] Amrashaa Singh, India's Courts Must Hold Social Media Platforms Accountable for Hate Speech, *Tech Policy Press*, April 11, 2025, *available at:* https://www.techpolicy.press/indias-courts-must-hold-social-media-platforms-accountable-for-hate-speech/ (last visited on May 5, 2025).
[3] *Ibid*.
[4] Vasudev Devadasan, "Conceptualising India's Safe Harbour in the Era of Platform Governance" 19 *Indian Journal for Law & Technology* (2024), a*vailable at:* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4727481 (last visited on May 5, 2025).
[5] "Beyond Safe Harbor: The Rise of Personal Liability in Platform Regulation" *SFLC.in* March 15, 2025, *available at:* https://sflc.in/beyond-safe-harbor-the-rise-of-personal-liability-in-platform-regulation (last visited on May 2, 2025).
[6] For details, see Article 19, Constitution of India, 1950.
[7] Matt Sheehan, "How Google took on China - and Lost" *MIT Technology Review*, December 19, 2018, *available at:* https://www.technologyreview.com/2018/12/19/138307/how-google-took-on-china-and-lost/ (last visited on May 3, 2025). (last visited on April 29, 2025).
[8] TOI World Desk, "Countries which have banned Telegram; From China, U.K. to Thailand" *The Times of India*, August 29, 2024, *available at:* https://timesofindia.indiatimes.com/world/countries-which-have-banned-telegram-from-china-u-k-to-thailand/articleshow/112891387.cms (last visited on May 3, 2025).
[9] Tom Phillips, "X goes offline in Brazil after Elon Musk's refusal to comply with local laws" *The Guardian*, August 31, 2024, *available at:* https://www.theguardian.com/technology/article/2024/aug/31/x-offline-brazil-elon-musk (last visited on May 3, 2025).
[10] Hamid Reza Saeidnia, et. Al, "Artificial intelligence in the battle against disinformation and misinformation: a systematic review of challenges and approaches" 67 *Knowledge and Information Systems* 3139–3158 (2025).

dealing with such issues of spreading disinformation are increasing liabilities of intermediaries and reducing the protection available through safe harbor provisions. For instance, Germany's Network Enforcement Act of 2017 was brought to control spread of misinformation. Facebook officials, including Mark Zuckerberg had to face investigation under this law for the failures of their company to handle hate speech issues.[11]

Research Objectives

- To examine effectiveness of the current legal framework related to safe harbor provisiona and content moderation in digital media in India
- To analyze international regime governing content moderation and digital media platforms' liability
- To provide recommendations to deal with the issues adopting a balanced approach.

Issues Related to Digital Media Regulation in India

The Indian government made comprehensive regulations regarding intermediaries and digital media in early 2021.[12] Ten global NGOs signed an open letter urging suspension of these regulations.[13] The government also previously fought with Twitter after the platform initially declined to delete tweets regarding some protests. The rules have especially impacted American tech giants such as Facebook, Twitter, and WhatsApp, requiring them to conform to governmental mandates of surveillance and censorship. Under the rules the intermediary companies should delete "illegal" content within three days of being notified. Social media platforms are obliged to provide user information on demand to the law enforcement agencies in certain situations. Encrypted messaging apps like WhatsApp should be able to track the "first originator" of messages and give this information to the government when requested under certain circumstances.[14] Some critiques allege that such rules are going to seriously alter the internet experience for ordinary Indian citizens by placing social media, news aggregation sites, music as video streaming platforms under direct government oversight. Indian officials say the rules are intended to check "abuse and misuse" of social media and empower users by instituting grievance redressal mechanisms.[15] WhatsApp's head Will Cathcart had resorted

[11] Rebecca Zipursky, "Nuts About NETZ: The Network Enforcement Act and Freedom of Expression" 42(4) *Fordham International Law Journal* 1325 (2019).
[12] The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021.
[13] Billy Perrigo, "India's New Internet Rules Are a Step Toward 'Digital Authoritarianism, 'Activists Say. Here's What They Will Mean" *Time*, March 12, 2021, *available at:* https://time.com/5946092/india-internet-rules-impact/ (last visited on May 1, 2025).
[14] Janjira Sombatpoonsiri and Sangeeta Mahapatra, "Regulation or Repression? Government Influence on Political Content Moderation in India and Thailand" *Carnegie India*, July 31, 2024, *available at:* https://carnegieindia.org/research/2024/07/india-thailand-social-media-moderation?lang=en&center=india (last visited on May 5, 2025).
[15] Billy Perrigo, "India's New Internet Rules Are a Step Toward 'Digital Authoritarianism, 'Activists Say. Here's What They Will Mean" *Time*, March 12, 2021, *available at:* https://time.com/5946092/india-internet-rules-impact/ (last visited on May 1, 2025).

to legal recourse against the rules necessitating violation of end-to-end encryption of the platform and had threatened to leave the country.[16]

Comparative Perspectives

A comparison of worldwide approaches, such as the United States' Section 230 of the Communications Decency Act, providing general immunity to internet sites, could shed light on achieving a balance between regulation and freedom. Comparing international approach in this regard with India's approach will highlights the need for well designed policies that protect users without stifling innovation or freedom speech and expression.

European Union

The Digital Services Act of 2022[17] under "Regulation on a Single Market for Digital Services" requires online platforms to identify and correct systemic risks in the spread of illegal content, disinformation, and threats to users' rights, e.g., freedom of expression.[18] While these steps attempt to tackle the international issue of disinformation, they may lead platforms to overact and remove legal content in an attempt to avoid fines. Such over-censorship is motivated by a "better safe than sorry" attitude whereby platforms prioritize respecting the rules over free speech.[19] The DSA proposes fines as high as 6% of companies' annual turnover in the event they fail to comply with the DSA's mandate. This financial burden has been criticised by stakeholders for being disproportionate. Such drastic penalties could drive platforms to shy away from providing free speech and invoke excessively restrictive content policies due to fear of legal liability. However, the DSA as it stands now makes platforms accountable without holding their representatives accountable as individuals.

United States

In the U.S., platform liability is perhaps most concerned about Section 230 of the Communications Decency Act of 1934. Section 230 shields platforms from liability for user-posted content, yet empowers them to

[16] Russell Brandon, "WhatsApp gives India an ultimatum on encryption" *Rest of World*, May 2, 2024, *available at:* https://restofworld.org/2024/exporter-whatsapp-encryption-india/ (last visited on May 5, 2025).

[17] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (Regulation - 2022/2065 - EN - DSA - EUR-Lex).

[18] "Your DSA Legal Representative in the EU", *available at:* https://prighter.com/product/dsa/?utm_source=google&utm_medium=cpc&utm_campaign=DSA-Rep-G-S-topical&utm_term=compliance&gad_source=1&gad_campaignid=21432281994&gbraid=0AAAAAC_lJIW-2Zc9Fm9XlrS_7Eyl1zDlZ&gclid=Cj0KCQjwlYHBBhD9ARIsALRu09r_WO8i5b48bVN-D6qUZqhu1KBzhoRTmiEzOEWxqWy7LuJGDu5WJsoaAhclEALw_wcB (last visited on May 11, 2025).

[19] Jacob Mchangama, et. Al, Thoughts on the DSA: Challenges, Ideas and the Way Forward through International Human Rights Law, *The Future of Free Speech, Justitia* (2022), *available at:* https://futurefreespeech.org/wp-content/uploads/2022/05/Report_thoughts-on-DSA.pdf (last visited on May 5, 2025).

moderate content in good faith. Section 230 has also been called the pillar of internet free speech because it allows platforms to host vastly diverse opinions without risking legal action. However, worries about spreading disinformation, hate speech, and abuse have been fueling growing demands for overhauls to Section 230. According to critics, the existing model allows sites evade accountability for the ill consequences of the content they host. Therefore, the critics demands that the law should be revisited to deal better with this issue.[20]

Brazil

In Brazil, the platform X (formerly-Twitter) was suspended and temporarily shut down after being ordered by a court for the absence of a legal representative in the country.[21] This is an instance of a trend in which governments are increasingly calling out platforms and their employees for not moderating content, particularly during elections and in the events of dissemination of political disinformation. Also, as per critics, Brazil employs personal liability as a tool for compliance by threatening to jail employees of platforms if platforms are noncompliant with takedown requests or refuse to generate user data.[22]

Russia

Russia has tightened its control over the virtual space by a combination of fines, requests for content takedown, and judicial action against domestic employees to compel compliance with its stringent online laws.[23] In recent times, Google, Twitter, and Facebook have come under sanctions and had to remove content that Russian governments deemed illegal. The ratification of the Sovereign Internet Law of 2019 has provided the administration with a tool to isolate Russia from the global internet, raising alarms over censorship and surveillance.[24]

[20] Ellen P. Goodman, Ryan Whittington, "Section 230 of the Communications Decency Act and the Future of Online Speech" *Rutgers Law School Research Paper* (2019), *available at:* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3458442 (last visited on May 5, 2025).

[21] Tom Phillips, "X goes offline in Brazil after Elon Musk's refusal to comply with local laws" *The Guardian*, August 31, 2024, *available at:* https://www.theguardian.com/technology/article/2024/aug/31/x-offline-brazil-elon-musk (last visited on May 3, 2025).

[22] Ibid.

[23] Alena Epifanova, "Deciphering Russia's "Sovereign Internet Law" 2 *DGAP Analysis German Council on Foreign Relations* (2020), *available at*: https://dgap.org/sites/default/files/article_pdfs/dgap-analyse_2-2020_epifanova_0.pdf (last visited on May 5, 2025).

[24] Erik Allerson, "Internet Censorship in Russia: The Sovereign Internet Laws and Russia's Obligations Under the European Convention on Human Rights" 31(1) *MINNESOTA JOURNAL OF INT'L LAW* 233 (2022).

China

China has never been relaxed in its grip on the virtual world, and websites have been required to abide by the Great Firewall a high-tech regime of content blocking and censorship.[25] International behemoths Google and Facebook have retreated from the Chinese market, citing the problems of doing business and maintaining global standards of free speech in China's censorship regime. Social media sites are required to employ staff to manually scan content, and failure to do this has implications such as fines, suspension of service, or cancellation of business licenses.[26]

Legal Structure of Indian Censorship

Indian Constitution envisages protection of freedom of speech and expression with "reasonable restrictions "under Article 19.[27] Judges have pulled on a narrow meaning for "public order" in order to approve net limits given the government's capacity to restrict access. India's regulatory strategy policy approach towards governing digital social media platforms is a combination of legislation and regulatory directives. The goal is to weigh the public morals, country's national security, and preservation of digital rights.

Amendment to the Information Technology Act, 2000

The key legislation that deals with online content in India is the Information Technology Act, 2000, Section 79, that provides "safe harbor" protection to intermediaries, excluding them from liability for third-party content on the condition that they follow due diligence requirements. The enforceability of content moderation policies under India's safe harbor law is a matter of law as well as ethics. Section 79 of the IT Act provides safe harbor to intermediaries against liability for content posted by users, if they do not intervene in inducing, modifying, or choosing content. The question that arises is whether content moderation, as a mandatory act of reviewing and affecting visibility of content comes within the 'safe harbor' provision or not.[28]The Act also mandates a fine for intermediaries or persons not cooperating with law enforcement agencies.[29] This provision imposes personal and criminal liability on individual employees, evidencing an excessive dependence on strict statutory controls to facilitate government compliance by private actors. Such measures can erode public trust in the online environment in the long term.

[25] Roya Ensafi, et. Al., "Analyzing the Great Firewall of China Over Space and Time" 1(1) *Proceedings on Privacy Enhancing Technologies* 61 (2015).
[26] "Beyond Safe Harbor: The Rise of Personal Liability in Platform Regulation" *SFLC.in* March 15, 2025, *available at:* https://sflc.in/beyond-safe-harbor-the-rise-of-personal-liability-in-platform-regulation (last visited on May 2, 2025).
[27] For details, see Article 19, Constitution of India, 1950.
[28] Vasudev Devadasan, "Conceptualising India's Safe Harbour in the Era of Platform Governance" 19 *Indian Journal for Law & Technology* (2024), *available at:* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4727481 (last visited on May 2, 2025).
[29] The Information Technology Act, 2000, Section 67C.

Significant amendments to the Information Technology Act, 2000 were proposed through the Jan Vishwas (Amendment of Provisions) Act, 2023. These were being introduced with a view towards ease of doing business as well as strengthening the law against cybercrimes. Five minor IT Act offenses were being deleted from the statute book in an attempt to lighten the load of the laws on business and help generate a business-friendly culture. Fines on serious offenses such as cyber fraud or stealing data were raised. This step aims to discourage criminals and motivate adherence to more rigorous cybersecurity standards. The amendments targeted intermediaries also by asking them to block hosting or transmitting unlawful content. Inability to follow these mandates now welcomes more stringent penalties to cast greater responsibility.

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

The IT Rules, 2021, were brought in to create a holistic regulatory regime for digital content in India. The rules place several obligations on intermediaries, such as social media platforms and OTT platforms, to make them accountable for the distribution of content.[30] According to the Rules, intermediaries are required to have a grievance officer. Additionally, intermediaries are classified as significant social media intermediaries; those with more than five million users are required to designate a resident grievance officer, a nodal contact person for round-the-clock coordination with law enforcement, and a chief compliance officer, all of whom must reside in India. The Chief Compliance Officer, being a senior staff member or managerial key employee in the organization, risks being imprisoned for a term of up to five years for offenses. The Rules also more specifically detail such duties as ordering the removal of content immediately after receiving government orders or court mandates and the posting of grievance officers.

Besides, there exists due diligence requirements. Intermediaries need to exercise due care by explicitly informing users regarding rules, regulations, privacy policies, and terms of use. They also need to take reasonable steps to avoid hosting, displaying, or transmitting unlawful or objectionable content. A three-tier mechanism has been put in place to address user complaints effectively. The mechanism consists of a self regulation by publishers as platforms are expected to resolve user complaints on their own.[31]Platform stakeholders constitute self-regulatory organizations which performs monitoring on their behalf. The oversight committee under Ministry of Electronics and Information Technology ensures compliance with the rules and adjudicates unresolved grievances. [32]OTT platforms are required to categorize content based on age suitability in categories like U (Universal), U/A 7+, U/A 13+, U/A 16+, and A (Adult).[33] So that children should not view any inappropriate content, platforms must impose parental controls as well as restrictions on

[30] Sakshi Pandey, "Constitutionality of the Information Technology Rules, 2021" *available at :* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4477004 (last visited on May 2, 2025).
[31] Sumeet Guha, Shreya Matilal, "INFORMATION TECHNOLOGY (INTERMEDIARY GUIDELINES AND DIGITAL MEDIA ETHICS CODE) RULES, 2021- A REASSESSMENT OF THE CONTOURS AND LIMITS" 8(2) *NUJS Journal of Regulatory Studies* 32 (2021).
[32] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, Rule 12.
[33] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, Code of Ethics.

access. Intermediaries must delete or disable access to unlawful content within 36 hours after being served a court order or government directive. The regulatory framework seeks to create an organized and responsible digital space while balancing censorship and protection issues for users.[34]

Digital Personal Data Protection Act, 2023

The "Digital Personal Data Protection Act" of 2023 aims at the privacy and personal data of Indian citizens. The act imposes stringent requirements on the digital platforms' storage, collection, and processing of personal data. The Act mandates certain categories of sensitive personal data to be stored and processed in India. The requirement enhances data security, implements Indian laws, and enhances the digital sovereignty of the nation. Digital platforms, or data fiduciaries, are required to obtain clear consent from individuals before collecting or processing their personal data. The Act also includes provisions for withdrawing consent and outlines the rights of data principals (data subjects). The Act gives special emphasis to safeguarding children's data. Platforms must implement effective age verification mechanisms and parental consent for users below 18 years of age to offer additional protection to children online. Organizational non-compliance with the Act's provisions is subjected to heavy financial sanctions in terms of turnover percentage, in order to render non-compliance unappealing.

Broadcasting Services (Regulation) Bill, 2024

Initiated in 2024, the Broadcasting Services (Regulation) Bill is aimed at putting in place one regulatory mechanism for various channels of broadcasting like conventional media, OTT, and digital news services.[35] This Bill is nothing short of replacing the outdated "Cable Television Networks (Regulation) Act" of 1995. The Bill proposes the creation of the Broadcasting Authority of India, an independent watchdog that oversees the broadcasting industry, upholds content standards, and settles complaints. The Bill prescribes a comprehensive content code with very detailed and specific guidelines on permissible content. Programmes, including OTT programmes, must be evaluated and certified by the Content Evaluation Committee before they are made available to the general public. For the purposes of enhancing national security and compliance with the law, the Bill requires broadcasting platforms to store user data only in India. While encouraging broadcasters to self-regulate, the Bill retains the government's right to interfere in cases of noncompliance. This ensures that content adheres to accepted standards without compromising the interest of the public. Both the "Digital Personal Data Protection Act" and the Broadcasting Services Bill highlight India's commitment to a secure, open, and citizen-centric digital space.[36]

[34] Naveen D. Chandavarkar, et. Al, "SWOC Analysis of the Information Technology

Telecommunications Act, 2023

The Telecommunications Act of 2023 updated India's telecommunication law by revising the Indian Telegraph Act of 1885. The Act provides a holistic framework for the regulation of telecommunication services, including digital communication platforms.[37] The definition of telecommunication services has been extended to include Over-The-Top communication platforms. The platforms are now subject to licensing and regulatory obligations akin to those applicable to traditional telecom operators. The Act empowers the government to suspend or limit telecommunication services or equipment for national security reasons. The Act also provides a framework for the interception and monitoring of communications under certain conditions. The Act has provisions for effective allotment and regulation of spectrum resources, enabling the use of sophisticated communication technologies and increasing connectivity. To give primacy to consumer interests, the Act enforces high levels of service quality and sets strong grievance redressal mechanisms for telecommunication services.[38]

These legislative changes are a testament to India's attempt to balance technological progress with strong regulatory control, providing a safe and consumer-centric digital environment.

Judicial Scrutiny

The case of '*Avnish Bajaj* v. *State (N.C.T. of Delhi)*'[39], brought into focus the limitations of the IT Act as it stood at the time, particularly regarding the liability of intermediaries. Avnish Bajaj, the head of Baazee.com which was an online shopping portal enabling the sale of products by commission-based transactions and advertising revenue, was arrested under Section 67 of the IT Act. Section 67 of the IT Act makes it a crime to publish or transmit obscene material. Further, charges under Sections 292 and 294 of the Indian Penal Code, which deal with the sale and dissemination of obscene material, were also brought. Bajaj's defense contended that the law makes a distinction between publishing obscene material and unintentionally making it available for listing. There was no proof to demonstrate that Bajaj or the website had transmitted or published the content.

The charge was due to an obscene video called "DPS Girl having fun" being on sale on the site. Baazee.com did not possess the goods on sale on its site, nor was it itself directly involved in payment and delivery, which were carried out by independent agencies. Bajaj's arguments pointed out that as soon as the company realized

the character of the video, it took corrective measures in 38 hours, even though there was an intervening weekend. The defense underlined that sites like Baazee.com cannot keep out misuse entirely but is required to act promptly once they are made aware of illegal material. This case highlighted the absence of provisions in the IT Act dealing with intermediary liability, something that was only filled later through amendments. The prosecution claimed that Bajaj had not stopped payments after realizing the video's illegal nature, and he has a role and responsibility in that regard. The probe did not prove payments were channeled through Baazee.com or that the video was directly accessible on the site.

The Delhi High Court observed that bail typically ought to be granted unless there is strong evidence to refuse it. The court noted that Bajaj cooperated fully with the investigation and no evidence of tampering was apparent. His non-Indian nationality did not preclude him from bail since he had close connections to India. The court also recognized arguments that denying bail in such matters would impact e-commerce development adversely in India. Nonetheless, it stated that economic concerns must not influence judicial rulings. The court stressed that while Baazee.com had taken steps to address the problem, intermediaries have to prove prompt and active diligence in such cases. It also pointed out the general implications of such cases for the developing digital economy and the desirability of clearer legislative guidelines. Section 67 IT Act: The court held that Section 67 read with Section 85 IT Act (which makes directors liable for company's acts) was applicable to Bajaj. This is because the IT Act recognizes deemed liability even if the company is not arraigned initially. The trial for Section 67 read with Section 85 IT Act against Bajaj was allowed to proceed. This case brought to fore the lack of proper legal definition and protection for intermediaries under the 2000 Act then. Though Bajaj's acts did not rise to the level of direct liability, the case brought out the difficulty for platforms in balancing compliance with innovation. The amendments later introduced to the IT Act brought in intermediary protections under Section 79, defining their obligations and liabilities in operating user-generated content.

In the path-breaking judgment of '*Shreya Singhal* v. *Union of India*'[40], the Supreme Court invalidated Section 66A of the IT Act as unconstitutional and clarified that only on receiving actual knowledge in a court order or government notification will intermediaries have an obligation to take down the content. This judgment reflected on the balancing of freedom of expression with delimiting the intermediary's duty.

Wikipedia platforms function through volunteer content moderation. Recent instances such as Indian court orders to have Wikipedia delete certain content or reveal editor identities reflect conflict between legal requirements and community-managed content handling.[41] These instances illustrate the problems platforms are facing in order to adhere to local laws while preserving their cooperative moderation regimes. Courts

must strike a balance between fundamental rights and technology regulation. Supreme Court can implement binding standards for content moderation. There are international precedents where action has been taken by courts against such platforms.

Challenges and Criticisms

In spite of legal provisions, various issues continue as far as national legal framework for digital media regulation is concerned. The 2021 IT Rules have been attacked for bestowing too much power on the government, which could result in censorship and the repression of dissent. This leads to government overreach and stringent censorship.[42]The constitution of a Grievance Appellate Committee with a majority of government-nominated members calls to question impartiality and the risk of arbitrary judgments. There also lies a lot of burden on intermediaries as platforms need to respond quickly to content takedown notices frequently without sufficient time for verification, risking over-censorship. Intermediaries are compelled to respond to removal requests within a very tight 72-hour window. This deadline could force platforms to be overly cautious, resulting in over-censorship and stifling lawful expression.[43]

There also lies a serious impact on privacy of individuals as requirement of tracing message originators threaten user privacy and encryption integrity. The obligation of intermediaries to determine the "first originator" of information presents some serious challenges to user privacy and end-to-end encryption integrity.[44] Platforms such as WhatsApp have objected to these provisions, claiming that they compromise user confidentiality and violate established privacy standards.[45]The IT Rules 2021 use vague and omnibus terms like "objectionable content," with no clear definition. This vagueness enables subjective interpretations, which can lead to unjustified censorship of content and suppression of free speech. The use of automated tools for content moderation raises issues, especially in a linguistically complex nation such as India. AI-based moderation systems tend to get caught up in contextual subtleties, resulting in the incorrect stripping away of legal content and the possible silencing of minority voices. The combined effect of the measures has caused adverse effect on free speech. Fear of punitive measures or removal of content has resulted in self-censorship by users, journalists, and content providers, subverting the democratic spirit of open debate.

Conclusion and Recommendations

The evolving Indian digital media landscape underlines the need to create an effective, well-balanced regulation that adheres to constitutional principles and caters to new issues. There lies a need to provide specific definitions of key terms within unlawful content to deter misusage. A good beginning point is providing definite, precise definitions of important concepts such as "unlawful content" to prevent their misuse and ensure consistent implementation.

Judicial review of notices of content removal is also required to safeguard against arbitrary or excessive government action. The government also needs to ensure judicial scrutiny of content takedown notices to guard against arbitrary actions. Policy-making needs to give precedence to open dialogue with civil society, industry stakeholders, and consumers of digital media. There should be an engagement of users in the formulation of policies to develop equitably balanced regulations.

Such cooperation would yield trust, legitimacy, and equity in the regulation process. Intermediaries would also have to be compelled to publish regular transparency reports detailing the number and nature of content takedown requests received and the resultant actions taken. The reports would bring accountability and facilitate public discussion on digital content regulation based on informed knowledge. As mainstream media credibility declines, social media sites have emerged as the crucial forums for public engagement and democratic discourse. Fading mainstream media credibility renders social media essential for public discussion. The technicalities which have been brought in by evolving technologies in creating, sharing, and regulating content bring technical skills into play among policymakers and regulators. Technical sophistication necessitates technical expertise. The judiciary will have to intervene actively to hold both digital platforms and the state accountable so that there can be a rights-based approach in conflicts involving digital media regulation.

Lastly, India must attempt to forge a middle ground that ensures freedom of expression, encourages responsible content practices, and encourages innovation in digital technologies. The legal system must evolve dynamically to engage with challenges of the modern day while ensuring that democratic values, transparency, and accountability are maintained in the digital age.