Page 5558 - 5566 **DOI:** *https://doij.org/10.10000/IJLMH.119605*

INTERNATIONAL JOURNAL OF LAW MANAGEMENT & HUMANITIES

[ISSN 2581-5369]

Volume 8 | Issue 2

2025

Digital Threats and Constitutional Rights: Regulating AI and Social Media in India

SHWETA[1]

ABSTRACT

Growing misuse of social media calls for an urgent move towards greater legal regulation to control the menace of fake news, hate speech and disinformation among democracies. Online interactions have undergone an evolution that threatens not only free speech, privacy, data protection, national security but democracy as a whole. Equally destructive is Artificial Intelligence (AI) threat to decisional and informational privacy. AI is the engine behind Big Data Analytics and the Internet of Things. While conferring some consumer benefit, their principal function at present is to capture personal information, create detailed behavioural profiles and sell us goods and agendas. Privacy, anonymity and autonomy are the main casualties of AI's ability to manipulate choices in economic and political decisions. The deployment of various AI systems has raised concerns about their potential negative impact on constitutional values enshrined in the Indian Constitution. In particular, the adoption of AI principles would have to strictly comply with the standards of anti-discrimination, privacy, the right to freedom of speech and expression, the right to assemble peaceably and the right to freedom of association as provided for in Part III of the Indian Constitution and interpreted by the Supreme Court of India. For instance, the right to privacy has been interpreted by the Supreme Court of India in the case of Justice K.S. Puttaswamy vs. Union of India to broadly include autonomy, choice, and control in the context of informational privacy. The subject matter assumes significance, in a democracy like India, which has notified a new regulatory regime - Information Technology Rules, 2021 and the Digital Personal Data Protection Act, 2023. The fact-finding review paper aims at mapping the evolution of laws governing online content in India. The study will be based on reviewing existing laws, regulations, policies, research papers, media reports and articles.

Keywords*: Social Media, Privacy, Artificial Intelligence, Intermediary.*

I. INTRODUCTION

In the present era, although online media platforms facilitate information sharing, they have also become conducive environments for the proliferation of illicit and dangerous content. The

[1] Author is a Research Scholar at Department of Laws, Panjab University, Chandigarh, India.

 [ISSN 2581-5369]

proliferation of unpleasant content in online spaces has proven detrimental. The circumspect exercise of free speech is also jeopardizing democracy. The emergence of contemporary technologies, like Artificial Intelligence (AI) and Machine Learning (ML), has intensified the issue to unprecedented levels. Field experts have expressed concerns over AI-enhanced algorithmic fairness, which results in discriminatory practices such as hate speech targeting race and gender, privacy violations, and user manipulation.[2]

The Hon'ble Supreme Court of India, in the case of Shreya Singhal vs. Union of India[3], stated that Parliament ought to enact new legislation to regulate social media. The Supreme Court has consistently urged the government to enact legislation that addresses the rapidly changing concerns posed by Information Communication Technology (ICT). Law enforcement agencies contest the social media service provider's attempt to classify itself as a "platform" instead of a "publisher", so allowing it to avoid accountability for the content shared on its sites. Print and electronic media are accountable for content published on their platforms, although internet platforms possess immunity. Likewise, there is a lack of transparency regarding the criteria for the removal of specific posts by service providers. The moderation and policies on collaboration with law enforcement do not consistently exhibit a uniform pattern. There was a pressing necessity for accountability for content publishing online. The Information Technology Rules, 2021 aim to address deficiencies in the regulation of the online space.

(A) Information Technology Act, 2000

The Parliament of India enacted the Information Technology Act, 2000[4] (IT Act), which became effective in October 2000. This marked a pivotal point in the development of digital media legislation in India, as it established the foundational legal framework addressing internet commerce and cybercrimes. The focus was primarily on addressing cybercrimes, despite the term 'cybercrime' not being defined in the Act and only addressing a limited number of computer-related offenses. The legislation was inadequately equipped to address challenges relating to social media and the internet. Critical matters like as free speech, privacy, data protection, and other essential elements pertaining to online conversation were not delineated. The Act did not delineate the parameters of due diligence for social media organizations. The Act revised multiple criminal and evidentiary statutes, notably the Indian Evidence Act, 1872 and the Indian Penal Code, 1860.

---

[2] Katarina Kertysova, "Artificial Intelligence and Disinformation: How AI Changes the Way Disinformation is Produced, Disseminated, and Can Be Countered", 29 *Security and Human Rights* 56 (2018).

[3] (2013) 12 SCC 73.

[4] https://www.indiacode.nic.in/bitstream/123456789/13116/1/it_act_2000_updated.pdf (last visited on October 03, 2024).

(B) Information Technology (Amendment) Act, 2008

The Amendment Act of 2008[5] implemented significant modifications to the existing IT Act of 2000, introducing explicit measures to expand the scope of cyber offenses covered by the original legislation. The new amendment introduced several clauses concerning individual data protection and privacy, along with measures to combat child pornography, voyeurism, and cyber terrorism in electronic and digital mediums.

To accommodate the swift expansion of social media in India, the Amendment included Section 66A, which addressed online offenses by penalizing the dissemination of "offensive messages." This provision was subsequently deemed unconstitutional by the Supreme Court in the case of Shreya Singhal vs. Union of India[6]. The court deemed it an unfair, disproportionate, and unreasonable limitation on the right to free speech. The Supreme Court reaffirmed that online communication should get the same constitutional protection of free expression as offline speech under Article 19 of the Constitution of India.

It introduced "government surveillance" through Section 69, which empowered authorities to intercept, monitor, or decrypt information. Websites can be blocked under this Section. Section 69B provides powers to collect traffic data from any computer resource. The aforementioned provisions are governed by various regulations, including the “Information Technology (Procedure and Safeguards for Interception, Monitoring and Decryption of Information) Rules, 2009” and the “Information Technology (Procedure and Safeguards for Blocking Access to Information by the Public) Rules, 2009,” which have been subject to governmental misuse.

Sections 69A and 69B collectively stipulate that governments may issue directives to ban websites for cybersecurity purposes, effectively constituting a form of internet censorship. These rules elicited considerable skepticism concerning the government's legitimacy to adopt measures that facilitate internet censorship.

The 2008 amendment explicitly defined "intermediary" to establish the responsibility for the offense. An intermediary refers to a social media platform that is a legal entity, which, on behalf of users, receives, stores, or transmits electronic records and offers services related to those records. It requires intermediaries to eliminate illegal content upon notification. The amendment mandates "due diligence" by these service providers when transmitting information.

Section 79 introduced by this amendment grants immunity to intermediaries about third-party

---

5 https://www.indiacode.nic.in/bitstream/123456789/15386/1/it_amendment_act2008.pdf (last visited on October 03, 2024).

6 *Supra* note 2.

content, data, or communication links uploaded by users, thereby absolving the service provider of liability for user-generated content.[7] This amendment introduced a significant change to the Act, which has become a prominent issue today as the government seeks to hold intermediaries accountable for users' content. Service providers, namely intermediaries, are lamenting the suppression of free speech under the guise of regulation.

(C) Information Technology (Intermediary Guidelines) Rules, 2011

To expand the parameters of "due diligence" by the intermediary, the government established thorough guidelines in April 2011. Nonetheless, intermediaries would find it exceedingly challenging to adhere to these standards due to the vast amount of content hosted and the intricate nature of legal infractions. The Supreme Court in the Shreya Singhal case[8] established two critical qualifications regarding the implementation of the clause. "Firstly, that the intermediary would only be obliged to remove or disable access to such content upon receiving actual knowledge that a court order had been issued directing it to do so". "Secondly, that the court order or the notification by the appropriate government authority must seek to restrict such content in conformity with reasonable restrictions laid down in Article 19(2) of the Constitution".

II. Overview of India's New Online Content Regulatory Regime

The Ministry of Electronics and Information Technology (MeitY) oversees the regulation of intermediaries associated with social media platforms, including Facebook, WhatsApp, Instagram, Twitter, LinkedIn, YouTube, and other applications. The Ministry of Information and Broadcasting (MIB) oversees the regulation of online news media and video streaming platforms, like Disney Hotstar, Amazon Prime, and Netflix, as well as traditional media such as print, radio, and television material.

India enacted the "Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021"[9] to primarily govern social media and expand its regulation to OTT platforms, online video streaming services, and digital news and current affairs providers. The new standards impose a set of obligations on global internet companies, necessitating greater accountability for the "misuse and abuse" of online platforms and the resolution of complaints

[7] The provision was in line with the US provision "Safe Harbour" law. Section 230 of the US Communication Decency Act, 1996 which is popularly known as the 'Safe Harbour' principles protect free speech along with the carriers of speech i.e. service providers.

[8] *Supra* note 2.

[9] https://www.meity.gov.in/writereaddata/files/Information%20Technology%20%28Intermediary%20Guidelines%20and%20Digital%20Media%20Ethics%20Code%29%20Rules%2C%202021%20%28updated%2006.04.2023%29-.pdf (last visited on October 03, 2024).

from individuals unjustly affected by hate speech and trolling. The Rules mandate social media platforms to disclose information regarding the "originator" of objectionable content upon request from the government or judicial authorities. The new regulations authorize the government to obstruct, remove, or alter published content or news within 24 hours. Additionally, Rule 4 of the 2021 Regulations stipulates a number of due diligence requirements for intermediaries in India. The "Due Diligence" clause has become excessively complex for intermediaries to manage. Noncompliance with these stringent standards would result in the forfeiture of the protections afforded to intermediaries under the "Safe Harbour" principle. Specifically, the personal accountability of intermediary officers for criminal and civil actions over third-party information. The provision requires the intermediary to appoint a grievance officer located in India, who shall be accountable for objectionable content.

The necessity to identify the originator of information directly conflicts with privacy rights, potentially undermining existing mechanisms that safeguard users' end-to-end encryption, which maintains the security of their messages.

The government has implemented the Code of Ethics for both online news media and video streaming entertainment platforms OTT (over the top media service). It requires digital news platforms to adhere to a set of codes of conduct that are outlined in the Cable Television Networks Regulation Act[10] and the Norms of Journalistic Conduct[11] of the Press Council of India. These codes of conduct serve as a general framework for the content that is published in print and television media. The Code of Ethics also mandates that video streaming platforms "exercise due caution and discretion" in light of India's multi-religious and multi-racial population. This is of paramount importance. Contents that are based on the beliefs, practices, or perspectives of any religious or racial group are highly sensitive, and it is imperative to exercise discretion while considering their sentiments. These provisions have been the subject of significant criticism from stakeholders, who are concerned about the potential for increased bureaucratic censorships and the suppression of artistic freedom in OTT content. India has recently experienced a surge in the number of court cases concerning the regulation of over-the-top (OTT) content. The grievances encompass a spectrum of emotions, including moral indignation toward sexuality depictions and wounded cultural and religious sentiments.

III. THE IMPACT OF ARTIFICIAL INTELLIGENCE ON SOCIAL MEDIA

[10] https://www.indiacode.nic.in/bitstream/123456789/15345/1/the_cable_television_networks_%28regulation%29.pdf (last visited on October 03, 2024).

[11] https://www.presscouncil.nic.in/ViewPdfContent.aspx?Page=DocumentsOfPCI&Title=Norms%20of%20Journalists%20Conduct,%202022 (last visited on October 03, 2024).

Social media platforms such as Facebook, Twitter, and Instagram have integrated AI technology into their algorithms to optimize the user experience.[12] AI is making waves in how it could change the way humans interact and consume content online. The most recent uses for AI include art, text and artificial voice generation.[13]

- **Advertising management.** AI-enabled tools help in advertising management and optimization. These tools can typically analyze and target ad variations, as well as perform customer segmentation.
- **Automatic posting and scheduling.** AI tools can integrate with social media platforms to schedule and post content at specified times or when they'd receive the most engagement.
- **Content generation.** Generative AI can be used to create social media posts with text or images or to create hashtags for a described post for an account.
- **Content moderation.** AI-enabled bots scrape data to find and filter out spam, guideline-breaking or inappropriate content. AI tools can also ban accounts that post that content.
- **Content recommendations.** AI tools can recommend video, text or image content that users may like based on previously consumed content.
- **Video filtering.** AI in facial recognition software helps recognize facial structures to identify users or overly edited filters on the user.[14]

(A) Risks of Artificial Intelligence in Social Media

a) AI bias is a significant concern, as it results in inherently biased decisions that are based on assumptions that are generated during the machine learning process. In the same vein, algorithms may be biased by biased humans.

b) Another issue is the promotion of echo chambers, in which thousands of users with a single opinion continue to share and reinforce a particular viewpoint or belief. Social media platforms have the potential to expose users to questionable content, such as posts that disseminate disinformation, by recommending content that users exhibit interest in. This has the potential to perpetuate users' preconceived notions.

c) Additionally, certain AI tools may acquire data on users that may be deemed

[12] Elsir Ali Saad Mohamed, *et.al*, "The Impact of Artificial Intelligence on Social Media Content" 20 *Journal of Social Sciences* 12 (2024).

[13] Alexander S. Gillis, "The impact of AI on social media", *TechTarget*, June 08, 2023, *available at*: https://www.techtarget.com/whatis/feature/The-impact-of-AI-on-social-media (last visited on October 02, 2024).

[14] *Ibid.*

intrusive by others. For instance, social media platforms may obtain information regarding a user's age, name, location, online activity, and photo metatags in order to generate a more precise advertising experience.

d) Additionally, there is apprehension regarding the dissemination of deepfakes[15] on social media platforms for the purpose of inciting malevolent social and political activity. A malicious actor could disseminate falsely generated images or recordings of another individual in order to achieve political or financial gain.[16]

IV. AI REGULATIONS IN INDIA

Currently, there are no specific codified laws, statutory rules or regulations in India that directly regulate AI.[17]

After a "deepfake" video clip of actor Rashmika Mandanna went viral on social media platforms in the year 2023, the Ministry of Electronics and Information Technology (MeitY) asked social media intermediaries to take such content down within 36 hours, a requirement outlined in the IT Rules, 2021.[18] The MeitY on March 01, 2024 issued an advisory[19] saying that all generative AI products, like large language models on the lines of ChatGPT and Google's Gemini, would have to be made available "with [the] explicit permission of the Government of India" if they are "under-testing/ unreliable". However, after the advisory came under criticism from experts for being ambiguous and vague, on March 15, 2024, MeitY issued a fresh advisory, dropping the requirement of obtaining "explicit permission" from the government. The latest advisory[20] said under-tested or unreliable AI products should be labelled with a disclaimer indicating that outputs generated by such products may be unreliable.

One plea has been filed by journalist Rajat Sharma in Delhi High Court against non-regulation of deepfake technology and has sought directions to block public access to applications and software enabling creation of such content. The other petition has been filed in Delhi High Court by Chaitanya Rohilla, a lawyer, against deepfakes and the unregulated use of artificial

[15] Deepfake technology is a type of artificial intelligence used to create convincing fake images, videos and audio recordings.

[16] *Supra* note 10.

[17] https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-india (last visited on October 03, 2024).

[18] Soibam Rocky Singh, "Stringent regulations could hinder growth of AI in India: experts" *The Hindu*, June 22, 2024.

[19] https://regmedia.co.uk/2024/03/04/meity_ai_advisory_1_march.pdf (last visited on October 03, 2024).

[20] https://www.meity.gov.in/writereaddata/files/Advisory%2015March%202024.pdf (last visited on October 03, 2024).

intelligence. The Delhi High Court on August 28, 2024, observed that deepfakes are going to be a serious menace in society and that the Centre should "start working on this".[21]

Nevertheless, various frameworks are being formulated to guide the regulation of AI, including:

- The National Strategy for Artificial Intelligence (June 2018)[22], which aims to establish a strong basis for future regulation of AI in India.
- The Principles for Responsible AI (February 2021)[23], which serve as India's roadmap for the creation of an ethical, responsible AI ecosystem across sectors.
- The Operationalizing Principles for Responsible AI (August 2021)[24], which emphasizes the need for regulatory and policy interventions, capacity building and incentivizing ethics by design with regards to AI. [25]
- Technology policy experts say that notification of the Digital Personal Data Protection (DPDP) Rules, which would provide teeth to enforce the Digital Personal Data Protection Act, 2023 notified in August last year, should be the top priority of the new government. The other key legislation that requires urgent attention is the Digital India Bill (DIB), which would replace the Information Technology (IT) Act of 2000 and revisit fundamental concepts like intermediary liability, fake news and deep fake content. The DIB is likely to regulate AI among other things.[26]

V. CONCLUSION

As noted above, there are currently no specific laws or regulations in India that directly regulate AI. As such, enforcement and penalties relating to creation, dissemination and/or use of AI are governed by related violations in non-AI legislation. Combating online disinformation is a gigantic problem as government regulations and steps taken by internet companies are not adequate. Regulatory mechanism tends to focus on content rather than addressing deeper structural obstacles that make it easy in disseminating false and misleading information. To address these issues, the government should have taken the legislative route to bring out such

[21] https://www.thehindu.com/news/cities/Delhi/start-working-against-deepfakes-delhi-hc-tells-centre/article68576870.ece (last visited on October 03, 2024).

[22] https://www.niti.gov.in/sites/default/files/2023-03/National-Strategy-for-Artificial-Intelligence.pdf (last visited on October 02, 2024).

[23] https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf (last visited on October 02, 2024).

[24] https://www.niti.gov.in/sites/default/files/2021-08/Part2-Responsible-AI-12082021.pdf (last visited on October 02, 2024).

[25] https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-india (last visited on October 02, 2024).

[26] Suraksha, "Data protection, AI rules top tech's wishlist for next govt", *The Economic Times*, June 03, 2024.

drastic provisions which may circumspect the right to privacy and free speech. The new IT Rules is a piece of subordinate legislation of the IT Act, 2000. No new law was enacted even after there is a dire need for such legislation.[27]

*****

[27] Ravi Shankar and Tabrez Ahmad, "Information Technology Laws: Mapping the Evolution and Impact of Social Media Regulation in India", 41 *Journal of Library & Information Technology* 295-301 (2021).

 [ISSN 2581-5369]