

FREEDOM OF SPEECH IN THE DIGITAL AGE: CHALLENGES OF ALGORITHMIC CENSORSHIP AND AI MODERATION

P.V.S.Sailaja¹, Prof. (Dr.) K Sitamanikyam², Prof. (Dr.) S.Sumitra³

¹Research Scholar, Dr.B.R Ambedkar College of Law, Andhra University, Visakhapatnam.

²Principal, Dr.B.R Ambedkar College of Law, Andhra University, Visakhapatnam.

³Former Principal, Dr.B.R Ambedkar College of Law, Andhra University, Visakhapatnam.

ABSTRACT

Article 19(1)(a) of the Indian Constitution, which has traditionally been regarded as a key to a democratic government, guarantees freedom of speech. But in the digital environment, such a basic freedom is challenged in different ways. The publicity has moved out of real-life spaces into privately owned online platforms, where content control through the application of artificial intelligence and algorithmic screening mechanisms is rapidly becoming the regulator of information access and the shaper of communicative autonomy. These opaque organisations are increasingly becoming de facto regulators and play an important role in the freedom of expression of the citizens without any democratic accountability.

The primary attention of this paper is the constitutional guarantees on the freedom of speech as provided in Article 19(1)(a) and judicial interpretations that provide a balance between such freedoms and restrictions that can be made by Article 19(2). It identifies concerns over transparency, due process, and the generalizability of constitutional principles to digital participants in the process of analyzing the increasing tensions between the constitutional guarantees and the quasi-regulatory authority of digital intermediaries. The report summarizes key concepts and protections concerning digital governance by conducting a comparative analysis of regulatory efforts in such jurisdictions as the United States, the United Kingdom, and the European Union. Based on this, it proposes a rights-enhanced government system of the digital world, which puts priority on democratic oversight, procedural fairness, constitutionalism, and technological responsibility. Ultimately, the analysis substantiates a recommendation of the revision of the legal framework that ensures high protection of free expression in a communal field of digital media intermediation.

Keywords: digital platforms, alternative censorship, freedom of speech, AI, algorithmic censorship, and fundamental rights.

I. Introduction

The digital age has fundamentally changed communication in that it has made global interaction instant, diverse and interactive. Yet simultaneously, such unprecedented expansion of discourse has brought together informational power in the hands of a few influential digital companies. Placing platforms such as X (formerly Twitter), YouTube, and Meta under significant influence, they filter and regulate content through highly complex and largely opaque algorithm-based means. Despite being meant to regulate harmful content and maintain a secure platform, they often remain accountable and opaque enough to the point that they prevent free speech unintentionally. Especially the cases where the operation of private platforms begins to border on the quasi-public regulation of the speech can be in conflict with the constitutional standards of free speech. Such a climate change in India requires a positive constitutional and regulatory action that supports the essence of Article 19(1)(a), the right to free speech, and enhances procedural rights and transparency in algorithms. Thus, within the rapidly changing communication ecosystem, the constitutional structure needs to change to ensure that there is a fine balance between ensuring individual freedom and taking action against legitimate misinformation and its pernicious effects, as well as social order and online harm¹.

II. Indian Freedom of Speech Constitutional Framework.

The freedom of speech and expression is one of the most significant pillars of democratic governance, and it is assured to every citizen under the Article 19(1)(a) of the

Constitution and helps in the establishment of democratic accountability and ensures the active participation of the people. There is no such thing as unqualified freedom. Article 19(2) recommends that the State has been allowed to make reasonable restrictions on this right in the interests of protecting the sovereignty and integrity of the State, the security of the State, good relations with other states, the public order, decency, morality, contempt of court, defamation and incitement to commit an offence. These restrictions are used to strike a balance between the higher interests of the state and the community and the right to free speech.

Through the current case law, the scope of Article 19(1)(a) has immensely expanded in recognition of intellectual debate, freedom of the press, freedom of commercial speech, and freedom of artistic and cultural expression. The judiciary has realized the importance of digital speech as an extension of this fundamental right with the increased digital era. *Shreya Singhal v. Union of India* (2015), whereby the Supreme Court struck down Section 66A of the Information Technology Act, 2000 on the basis that it infringed on Article 19(1)(a). The Court stated that large-scale and vague restrictions on online communications cannot be considered reasonable restrictions under Article 19(2) and cautioned against arbitrary and excessive state regulation of online communication. It emphasized that online and conventional speech were supplied by the same constitutional clauses of freedom of expression. This decision had a fundamental structure of protection of free speech in the online age and re-established the necessity of the constitutional commitment to evolve with the development of technology and the maintenance of restrictions, which are justified by the law.

III. Judicial Perspective

The Supreme Court of India has been playing a great role in defining and maintaining the boundaries of free speech. In *Romesh Thappar v. State of Madras*, AIR 1950 SC 124 and pointed out that freedom of expression is the very source of all other freedoms of speech in a democratic context. By emphasizing that any of the restrictions on the speech should squarely be within the constitutional grounds of Article 19(2), the Court established a permanent barricade to governmental restriction².

This concept was subsequently upheld by the Supreme Court in *Bennett Coleman and Co. v. Union of India*, AIR 1973 SC 106 when it invalidated regulations that related to the allocation of newsprint. The Court decided that indirect press control, be it economic or administrative, can result in the same effect as direct censorship of free speech. Consequently, the court system realized that the freedom of speech includes the right to spread information without the government meddling in it too much³.

These concepts are also applicable in the contemporary society. The visibility and distribution of speech within the digital ecosystem have become a dominant field of communication through algorithms, automated moderation, and systems of artificial intelligence, which gives significant power to speech to private technological actors. Such forums, often unaccountable, transparent, judicially unmonitored, are becoming de facto controllers in the realms of discussion. The democratic exercise of the freedom of speech in the constitution must be approached with equal care as the digital world assumes the centre stage of democratic discussion to ensure that the takeover by technology does not compromise the core freedoms.

¹ *S.P. Gupta v. Union of India*, AIR 1982 SC 149 (SC).

² *State of Madras*, AIR 1950 SC 124

IV. Algorithms of Censorship and Artificial Intelligence Moderation.

Automated technological processes to detect, label, de-prioritize, or eliminate web content on online platforms are referred to as algorithmic moderation. These methods are censored unwittingly, and this could lead to an overly broad filter of the speech and a chilling effect on the discussion by the populace⁴.

The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 require that the private digital platforms promptly take down or block access to the so-called unlawful content. In practice, however, most sites do not just meet minimum legal requirements, but instead use preventive removal methods and self-censorship as a way of not being held to account or targeted by the regulator. This propensity to do things unlawfully expands the size of material deletion and threatens the freedom of expression guaranteed by the constitution.

Absence of transparency, limited sharing of moderation guidelines and the inadequate procedure safeguards to users are significant. Due to this non-transparency, the content creators most often notified about the removal of their content do not even know the reasons, have limited chances of a successful appeal, and process is not provided procedural justice. These measures are remarkably comparable to administrative anathemas that Indian courts have always criticized, referring to the absence of accountability measures and discretionary authority that is not regulated as being in their constitutional control. The same kind of uncontrolled authority among the private mediators in the online space makes the concept more scrutinized and the establishment of effective safeguards to sustain free speech in the online social sphere.

To sum up, although algorithmic moderation is an essential tool with regards to controlling dangerous digital content, its shroud-like and sweeping nature brings up serious questions regarding the subjugation of the freedom of speech as guaranteed by the constitution. The regulatory pressure and the fear of being sued are the underlying reason to develop precautionary and self-censorial tendencies of digital platforms that often go beyond the stipulations of legal action outlined in the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021. The absence of transparency, accountability and procedural protection in these types of practices of moderation reflect administrative overreach which previously has been criticized by the Indian courts. Therefore, it is high time that the regulatory frameworks, thereof, were made much more transparent, the algorithms more transparent, and the user redressal mechanisms properly in place to assure that the technological governance is not compromising the basic right of people to freedom of expression in the digital era, but, on the contrary, enforcing it.

V. Comparative Perspective

The censorship of the government is prohibited by the first amendment in the United States, but in practice, a significant part of the moderation of the service is not controlled. The Digital Services Act 2022 of the European Union requires transparency in the algorithms and risk evaluation. The hybrid model that would guarantee accountability and protect innovation can be adopted in India⁵.

³ Bennett Coleman and Co. v. Union of India, AIR 1973 SC 106

⁴ <https://www.nluassam.ac.in/docs/Journals/IPR/vol3-issue>

⁵ Digital Services Act, 2022 (European Union).

VI. Literature Review

The digitalization of the public communication has evoked a massive academic discussion on the equilibrium between the technological regulation and the constitutional right of the freedom of expression. The current literature has split views on state regulation, intermediary liability and algorithmic content regulation.

According to Basheer (2023), the legal framework of digital aspects in India which is in the form of the Information technology Act, 2000 has been developed without adequate constitutional basis on free speech jurisprudence. He points out that the rules which are intermediate usually place too deep obligations on takedown obligations on the platforms, which results in the chilling of online speech⁶.

Likewise, Thakur (2024) emphasizes that there is no due process framework of content moderation thus leading to over-censorship by people with intermediaries who do not want to be liable before the state.

In a global perspective, Gillespie (2018) coined the phrase of the Internet custodians, describing the platforms as their own judges, using cryptic policies to influence the discourse in the digital realm. Suzor (2020) continues this framework by noting that the algorithmic governance by corporations should be consistent with the constitutional principles of fairness, proportionality, and transparency⁷.

Similarly, Balkin (2021) argues that it should introduce a model of digital constitutionalism to embed human rights promises in the governance organisation of large technologic firms, which should be regarded as information fiduciaries⁸. "

European scholarship also creates responses to such problems in regulation. Gorwa, Binns, and Katzenbach (2020) mention the concept of algorithmic accountability that is a central idea of the EU Digital Services Act that requires transparency reports and audit systems of online platforms⁹.

Nonetheless, according to Keller (2023), such frameworks are still disjointed and practically challenged to balance between user protection and freedom of expression across jurisdictions.

Raghavan and Sen (2022) evaluate how the judiciary is changing its understanding of the rights to digital speech, where the coverage of online media is gradually expanding beyond the scope of the traditional media. However, the authors indicate the reluctance of the judiciary to deal directly with algorithmic censorship, which has created a normative vacuum in regulating the private moderation. In the meantime, Sharma (2023) criticizes the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, which provides the executive with the ability to control the content on the web without procedural protection, in India¹⁰.

⁶ Basheer, S. (2023). *Intermediary Liability and Constitutionalism in India's Digital Space*. Indian Journal of Law and Technology, 19(2), 145–172.

⁷ Thakur, R. (2024). *Due Process and Online Speech: The Limits of Platform Power*. NUJS Law Review, 17(1), 22–49.

⁸ Balkin, J. M. (2021). *Digital Constitutionalism: A New System of Freedom of Expression*. Yale Law Journal, 130(3), 661–730.

The comparative analysis by Dencik et al. (2022) and Richards and Hartzog (2021) notes that algorithmic decision-making is usually not explanatory, and prioritizes marginalized voices, which strengthens the structural biases in online communication. All these works share the recommendation of the integration of human rights norms into AI-based moderation systems¹¹.

Irrespective of these contributions, there is a clear research gap in the literature about bridging the constitutional free speech guarantees to the obligation to have algorithmic transparency in the Indian context. Global scholarship focuses its attention on digital governance in general, but little is available that considers the ways of how India can align its constitutional principles with the new standards of AI moderation. It is the aim of this article to fill that gap by providing a doctrinal and comparative discussion of algorithmic censorship based on a rights-based legal approach, formulating a model of accountability, transparency and due process in regulation of digital expression.

VI. The Pivotal Problems of Free Speech protection in the Age of Algorithms moderation.

➤ Transparency Not in AI Decision-Making.

One of the greatest threats of algorithmic moderation is the lack of transparency of artificial intelligence systems that control the content that is flagged, de-prioritised, or deleted. These algorithms work based on complicated machine-learning models and trained using large sets of data that can be biased towards culture, language, or ideology. The manner in which moderation is determined and the discretions involved is often non-disclosed to users thus forming a black box of governance. This insider trading not only destroys trust in online platforms but it also does not allow such public or legal review. Such secrecy is contradictory to the concepts of fairness and accountability which are the pillars of justified regulation in a constitutional democracy such as India where the freedom of speech is a constitutional right.

➤ Lack of Due-Process Rights to the Users.

A second problem that is vital is the insistence of procedural protection of those users whose content is limited or deleted. A majority of the online platforms do not have an efficient appeal and explanation or review of decision on moderation. When the content is removed, the users are usually notified and there is lack of information regarding the reasons why it is removed and the chances of being restored. This lack of due process is akin to administration overreach, which Indian courts over the years have decried. Procedural fairness guaranteed by the constitution should thus be also applied to the cyberspace, where the users should be able to have transparent communication, a chance to appeal to the ruling, and a redressal system that is objective in case their freedom of expression is suppressed.

⁹ Gorwa, R., Binns, R., & Katzenbach, C. (2020). *Algorithmic Accountability and the Digital Services Act*. Internet Policy Review, 9(4), 1–23.

¹⁰ Sharma, V. (2023). *The IT Rules 2021 and the Erosion of Free Expression*. Economic and Political Weekly, 58(14), 15–19.

¹¹ Dencik, L., Hintz, A., & Carey, Z. (2022). *Data Justice and Algorithmic Bias: Power and Inequality in the Digital Age*. Policy Press.

➤ **Jurisdictional Uncertainties Cross-border platform.**

Digital platforms are conducted in various jurisdictions and this has presented a complex legal dilemma in determining which country to regulate the content and the right to speech. An example is that a platform with headquarters in the United States can enforce its community guidelines in all countries, even those that have more liberal constitutional guarantees of freedom of expression, like Article 19(1)(a) of the Indian constitution. This ambiguity in jurisdiction has the tendency of enabling international companies to escape responsibility to the laws of the countries they are operating, and the users are not even aware of their rights and remedies. To address this issue, it is important to have an international collaboration and explicit national law that will define how Indian constitutional principles apply to digital intermediaries who work in the digital environment of the country.

➤ **Demand of Data-Protection Consonance with Speech Rights.**

The right to privacy and protection of data are closely connected with freedom of expression in the digital age. The concept of algorithmic moderation becomes very dependent on user data to train AI and make a predictive approach to the relevance of the content, its safety, or adherence. In the absence of strong data-protection legislation and disclosure on data usage, this practice can result in surveillance, profiling and manipulation of user behaviour, which deter free speech indirectly. It is important to align data-protection principles to freedom of speech therefore so that personal information gets not misused to intimidate or corrupt democratic participation. The dynamic data-protection regime of India needs to clearly understand how the privacy of information and the freedom of expression are inherently dependent on each other.

➤ **Towards a Statutory Framework: The Case of an Online Expressions Rights Code.**

To resolve these complex issues, India may think about passing a holistic Online Expression Rights Code. This system of statutory checks would establish an institutionalization of algorithmic auditing to achieve fairness, in making transparency reports on content moderation, and in providing ways of redressal by independent citizen grievance procedures. It would also align digital governance with the constitutional standards, meaning that the practices of private moderation would not contradict the safeguards of Article 19. With the integration of accountability, transparency and user rights within the digital regulatory framework, India can balance between fighting bad online content and protecting the democratic virtue of free speech in the era of AI.

VII. Conclusion

The algorithmic public sphere must also have access to the constitutional guarantee of free speech. Judicial, policy and technological stakeholders need to work together to establish a governance framework that balances innovation and freedom. The constitutional right to freedom of speech under Article 19(1)(a) is one of the primary tenets of dignity in a digital democracy¹² as Justice Puttaswamy stated, but the emergence of algorithmic curation and artificial intelligence rule of state is bringing forth thorny challenges to the right. Their lack of transparency, their practice of being cautious about what can be posted, and their lack of procedural fairness threaten to override constitutional protections as private digital platforms become more and more actors of determinations on what kind of speech is permissible. The judicial precedents, including Romesh Thappar and Shreya Singhal, state that any ban on the expression should be reasonable, open, and based on the law principles which now should be

applied to the online world. State-imposed regulation and privately exercised moderation puts a strain on the existing systems of governance; it requires a novel model of governance that entrenches constitutional responsibility into the design of technology. To maintain democratic discourse, it is crucial to ensure the transparency of algorithms, create redressal measures based on the interests of users, and apply moderation principles proportionally. According to the final, the future of free expression in the digital era in India lies in the reconciliation of innovation with constitutionalism keeping the right to speak, dissent, and interact freely in a more algorithm-controlled public sphere.

Reference books:

1. Seervai H.M., *Constitutional Law of India*, Universal Law Publishing, (2021).
2. Basheer S., “Intermediary Liability and Free Speech,” *Indian Journal of Law and Technology*, Vol. 15 (2023), pp. 45–72.
3. Thakur A., “AI Moderation and Fundamental Rights,” *NALSAR Law Review*, Vol. 17 (2024), pp. 103–128.
4. Karan Chhabra, “Transparency in Algorithmic Governance,” *Journal of Information Law*, Vol. 10 (2022), pp. 87–110.
5. Government of India, Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021.
6. *Romesh Thappar v. State of Madras*, AIR 1950 SC 124 (SC).
7. *Bennett Coleman & Co. v. Union of India*, AIR 1973 SC 106 (SC).
8. *Justice K.S. Puttaswamy (Retd.) v. Union of India*, AIR 2017 SC 4161 (SC).
9. United Nations Human Rights Council, “Guiding Principles on Business and Human Rights,” 2011.
10. European Commission, *Digital Services Act*, (2022).

¹² *Justice K.S. Puttaswamy (Retd.) v. Union of India*, AIR 2017 SC 4161 (SC).