
LEGAL IMPLICATIONS OF SOCIAL MEDIA SHADOW BANNING AND CONTENT DEMOTION: AN ANALYTICAL STUDY

Ankit Kohli¹ & Amber Srivastava, Assistant Prof.²

ABSTRACT

Social media has been a game-changer for communication, democratic participation and dissemination of information. Platforms like Meta Platforms, previously known as Facebook, X Corp. (formerly Twitter), and Google LLC have become the custodians of online visibility, with algorithmic governance. Issues surrounding moderation that are seen as controversial involve shadow banning and content demotion, in which users' material is limited without explicit notification. As a countermeasure to misinformation, hate speech and bad content, platforms argue that such measures are essential, but they create a multitude of problems concerning transparency, procedural fairness, freedom of expression, and accountability. The present article analytically explores the constitutional and compositional considerations, intermediary liability regime principles and principles of digital governance in the case of shadow banning. The present article would analytically debate the constitutional and compositional considerations, while examining the principles of intermediary liability and digital governance in the case of shadow banning, especially in the Indian context. It examines whether algorithmic suppression is indirect censorship, and examines the public/private balance of autonomy of the platform and fundamental rights of users.

Keywords: Social Media Regulation, Shadow Banning, Content Demotion, Freedom of Speech, Algorithmic Governance, Intermediary Liability, Digital Rights, Constitutional Law, Online Censorship, Platform Accountability.

¹ Law Student, Law college Dehradun, Uttarakhand University, Dehradun

² Assistant Professor, Law college Dehradun, Uttarakhand University, Dehradun

Introduction

Today, communication systems have undergone a completely new development in the service of the 21st century with digital platforms and social network technologies. Social media has become a great vehicle for political, economic, public discourse and journalistic activity rather than just communication within the user's social network. Rather, social platforms like Meta Platforms, X Corp. and Google LLC serve as key entry points by which people interact with news, share opinion, rally social movements and interact in democratic dialogue. The existing media institutions are different because they have identifiable editorial policies, are regulated by statute and have factual regulations. Social media operate in a different way, their visibility, reach and engagement are largely determined by algorithmic systems. The move from managing it with human editors to governing it with algorithms has brought fresh legal and constitutional issues.

Practices that have become most controversial are shadow banners, which are a moderation technique used by many algorithms that basically hide the content of a specific user and keep them from knowing about it. Viewing content demotion closely, posts' rank is lowered in their visibility by algorithm without actual deletion. These practices are not like usual censorships as users are usually not aware whether any restrictions are imposed on them. Platforms advocate for such actions as tools for tackling misinformation, hate speech and extremism, and harmful online activity, but critics say shadow banning functions as invisible barriers on free expression, and it threatens democratic transparency.

The legislation's difficulties are because social media businesses are hybrid entities. They are companies that operate their platforms as contractors, while on the other hand they carry out quasi-public functions by hosting massive public debate on the platforms. In a democracy, such as that of India, freedom of speech and freedom of expression is a fundamental right protected by Article 19(1)(a) of the Constitution. The protections which are traditionally enshrined in the Constitution are protections against the State, however, it is not clear how much protection might extend to some private digital platforms that may have a potential to shape the fundamental freedoms. Increasingly political discourse is taking place online, and algorithmic decisions reached by private institutions can have a significant impact on political participation and social reputation, as well as on professional opportunities.

In addition, because of the lack of transparency, there are questions about a non-discriminatory

process, natural justice, and accountability. Users are typically unaware of what constitutes moderation of acceptable or acceptable practices, are unable to effectively challenge controls on things that should be visible, and may have weak access to grievance procedures. There is no transparency to platform governance — speech can be suppressed without being formally banned. Thus, shadow banning and content demotion not only raise technological questions, but constitutional and human rights issues that warrant immediate study and discussion by scholars and lawyers.

The present analytical attempt therefore aims at reflecting on the legal implications of shadow banning in the field of intermediary liability, constitution guarantees, algorithmic accountability and the emerging global models of technology governance. The article argues on the basis of judicial precedent, the evolution of legislation and theoretical discourse that algorithm-based practices of moderation can be considered valid content governance or privatized forms of censorship in the digital era.

Main Discussion

As the intermediary liability landscape continues to evolve, constitutional safeguards and controls exercised over algorithms must be considered when analysing the legal implications of shadow banning and content demotion. What is happening on social media, is not really a publishing of media, it rather bridges the space between platforms and to a certain degree user-generated content. Social web sites are primarily spaces with user-generated content, which, in turn, is platforms and not publishers. Under the Information Technology Act, 2000, the intermediaries will remain immune until they perform the task of 'due care' which would involve removing the material that is prohibited if and when they get a legitimate order. But shadow banning does not occur within the process of official regulation because it's not added to the list of banned content and is not deleted from shoppers' feeds; instead, it simply becomes invisible because it is hidden from viewers' feeds through mechanisms of automatic algorithms.

This occurs and leads to the phenomenon of algorithmic censorship or algorithmic control over speech, control via design, rather than by law. Recommended systems, such as those developed by Meta Platforms, X Corp, Google LLC and others, employ algorithms to show the most clicked or watched content, prioritize community standards or behaviour, or account for business factors. These amounts are shared with the public, meaning that users might not be

aware of their speech being altered, thereby reducing transparency and fairness.

In *Shreya Singhal v. Union of India*, the constitutional importance of the invisible restrictions comes to the fore. The landmark Supreme Court of India ruling in this case struck down Section 66A of the Information Technology Act as "chilling" on the right to freedom of speech, due to the ambiguity and over-breadth of it. The requirement on restrictions on expression is rooted in the need for clear, narrowly-tailed and procedural safeguards, the Court said. The reasons given for the case of intervening in freedom of expression are relevant and are also very apt for algorithmic censorship. Shadow banning does work the same: there's no clarity for people around what they can and can't post, there is no transparency around why their speech is being banned. It can be possible that it "chills" lawful speech and expression, similar to the Court in *Shreya Singhal*.

Further, the court in its judgment had stated that these intermediaries were taking the back seat from the court and doing so without any kind of guidance from the court or government. Platforms can be involved (alone) in decisions which promote and demote content without any clearly documented and accountable basis, they argue, they are exercising quasiadjudicatory powers, which affect the right to speech. So, the question is, are there grounds for a constitutional protection of algorithms used by the private sphere in public communication functions?

What is important here, is particularly procedural justice and natural justice. The standards of traditional law are that anything that affects rights would be in accordance with a fair procedure – one involving notification and opportunities to contest the action. All of these protections are rendered invisible governance because of Shadow banning. This practice might therefore clash with the general constitutional recognition in *Maneka Gandhi case v. Union of India* of the need to ensure restrictions on liberty must be just, fair and reasonable.

The issue of algorithmic responsibility was turned into an international one in *Gonzalez v. Google LLC*, leading to doubt that algorithmic recommendations cause content dissemination. That there is no passive conduit, there is a system of decisions, which may affect public discussion, the case said. Content demotion and shadow banning, in turn, are the examples where algorithmic architecture may contribute – without officially banning the speech/content – creating a shaping of the discourse.

Social media can be important communication infrastructures in democratic societies in particular for politics. In various countries, both global and in Europe, there is ample evidence of how social media has been identified and characterised as a new form of the public forum, like in *Packingham v. North Carolina*. The affect of limited sight visibility – even the opinion level of a particular individual – can have significant implications for democratic involvement. This needs to be accompanied by a set of accountability measures to ensure no arbitrary and discriminatory action by digital platforms against users as they play an increasingly important part in meeting their needs.

The question of the cutting edge of the regulation is a practical question which strikes the balance between innovation and private citizen's rights and constitutional rights. If it's either or, digital regulation may become government censorship, and digital unregulated algorithmic authority will certainly become 'privatisation of freedom of speech' if the government's involvement is too intrusively large. As a result, transparency specifications, algorithm audits and effective grievance redressal mechanisms are now part of the conversation in today's legal landscape, and have been recommended to moderate practices to line up with constitutional values.

Finally (but most importantly), know the different ways content they are about to use is individually moderated – shadow banning and demotion. Now rather than legislative or judicial government the corporate algorithm government is in control of some of the tail. As we move into this new era of digital governance, the clear, proportional and "chilling" principles of the law of the Constitution are important considerations. Ensuring adherence to these principles, one of the biggest legal challenges of today, is "algorithmic moderation".

Case Laws

The law from complaints about digital speech, intermediary liability, and platform responsibility is instructive of issues that have been discussed regarding shadow banning and demotion of content.

In *Shreya Singhal v. Union of India* the Supreme Court of India held that Section 66A of the Information Technology Act, 2000 is vague and restrictive on freedom of speech. The Court highlighted that the right to freedom of expression is subject to the Constitution requirement for restriction to meet the constitutional test of reasonableness and legality. The judgment is

also relevant to the shadow banning of debates, as invisible restrictions on speech - for whatever reason - could have and have the same effect as a chilling effect with no clear legal basis.

In *Faheema Shirin v. State of Kerala*, the Court in Kerala took a holistic views of the right to education and freedom of expression enshrined under Article 21 and thus uplifted access to internet to be an inseparable part of the right to education and freedom of expression. When made, it will reinforce the claim that the use of the internet and online platforms is a contemporary aspect of constitutional freedoms.

In the *Packingham v. North Carolina* case, the United States Supreme Court recognized the importance of social media platforms, calling them the “modern public square.” There are arguments for the restriction on access to be significant for fundamental speech rights.

The case filed by the family of the 2011 killed terrorism victim that they're claiming is *Gonzalez v. Google LLC* in which Google LLC had their blog recommending extremist content. The case raised the question about the level of passivity or activity of the recommendations – whether they were an active removal procedure or an active indication of the existence of such a procedure? Case asks is an assisted removal of content from the platform—an “active” or “passive” participation under Section 230 of the U.S. Communications Decency Act? The Court ended up stitching up liability and merely upholding the status quo in Section 230, the ruling ultimately led to a worldwide discussion on algorithmic accountability. It is relevant to point out that recommendation algorithms are not neutral algorithms, but instruments along with their demotion counterparts which must affect the visibility of speech in the Company's context of shadow banning.

Moreover, what *Maneka Gandhi v. Union of India* holds is that there must be fair, just and reasonable procedure to anyone that is to take such an action which has impact on the personal liberty. Algorithmic suppression without notice, explanation or remedy can undermine basic tenets of procedural fairness that form the basis of constitutional government, if applied by analogy.

Suggestions

There are a few changes that could be made to overcome the legal repercussions of shadow

banning and content demotion.

First, platforms have a responsibility towards consumers to implement a mandatory transparency standard, which means they must disclose that a video is invisible if it is visible, and disclose that it is not if it is not visible. Users have to be notified when their posts get removed, demoted, or when it is limited when the algorithm is used.

Second, procedures should be established which are comparable to natural justice principles. A description of the reasons for moderation actions and guidelines on appeal processes need to be provided to a user.

Third, autonomous regulatory entities or digital ombudsman can hold accountable in the absence of too much state involvement. Such entities enable an audit of algorithms for bias, discrimination or political manipulation.

Fourth, there are in the legislation the ability to set the concept of obligations of intermediaries but not to hinder innovation and freedom of enterprise. A good governance will be obtained by the government civil society platform, balance.

Finally, digital literacy initiatives should ensure that offering data on how algorithms work would enable users to engage in online environments and inquire about them.

Conclusion

The governance of shadow banning and Content demotion is newfound in a digital vs. legal relationship. Combating harmful content through social media is a critical tool, however there is a lack of transparency around the algorithm and as such a risk of lack of freedom of expression, democratic participation and procedural justice. The rules should be what it is now: there is less of a control over the public discourse by techno-less means.

A balance between competing interests, platform autonomy and state regulation, and individual interests will be key in working out online speech regulation. The preservation of a free space in the network, rather than censorship, is important when building regulatory frameworks that are grounded in rights, rules and procedures with an incomplete sense of "transparency" and foster autonomy and accountability of algorithms. The concept of "shadow ban" and practise is a more complex constitutional issue for the law scholars, as well as for the policy makers, and in the algorithmic age of power.

References

Statutes

- Constitution of India, 1950 – Article 19 and 21.
- Information Technology Act, 2000.
- Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021.
- European Union Digital Services Act, 2022.

Cases

- **Shreya Singhal v. Union of India**, (2015) 5 SCC 1.
- **Faheema Shirin v. State of Kerala**, 2019 SCC OnLine Ker 1733.
- **Packingham v. North Carolina**, 582 U.S. (2017).
- **Maneka Gandhi v. Union of India**, (1978) 1 SCC 248.
- **Gonzalez v. Google LLC** , 598 U.S. (2023).

Books

- **Lawrence Lessig**, *Code and Other Laws of Cyberspace*, Basic Books.
- **Tarleton Gillespie**, *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*, Yale University Press, 2018.

Journal and Articles

- **Jack Balkin**, “Free Speech in the Algorithmic Society ,” *Yale Law Journal*.