



# THE EVOLUTION OF HUMAN RIGHTS IN THE DIGITAL AGE: ADDRESSING ONLINE HATE SPEECH AND FREEDOM OF EXPRESSION

**Pallavi Chaurasia**  
Research Scholar,  
Vikrant University, Gwalior,  
PHD Law

---

**DECLARATION:** I AS AN AUTHOR OF THIS PAPER /ARTICLE, HERE BY DECLARE THAT THE PAPER SUBMITTED BY ME FOR PUBLICATION IN THE JOURNAL IS COMPLETELY MY OWN GENUINE PAPER. IF ANY ISSUE REGARDING COPYRIGHT/PATENT/OTHER REAL AUTHOR ARISES, THE PUBLISHER WILL NOT BE LEGALLY RESPONSIBLE. IF ANY OF SUCH MATTERS OCCUR PUBLISHER MAY REMOVE MY CONTENT FROM THE JOURNAL WEBSITE. FOR THE REASON OF CONTENT AMENDMENT /OR ANY TECHNICAL ISSUE WITH NO VISIBILITY ON WEBSITE /UPDATES, I HAVE RESUBMITTED THIS PAPER FOR THE PUBLICATION.FOR ANY PUBLICATION MATTERS OR ANY INFORMATION INTENTIONALLY HIDDEN BY ME OR OTHERWISE, I SHALL BE LEGALLY RESPONSIBLE. (COMPLETE DECLARATION OF THE AUTHOR AT THE LAST PAGE OF THIS PAPER/ARTICLE

## ABSTRACT

*The online age has changed the debate on the human rights, posing new questions regarding the freedom of expression and cyber hate speech control. Despite the fact that the internet and the social media have democratized communication, the rate at which harmful and discriminative content that negates equality, dignity, and social order spread has been high. The present paper studies the evolution of human rights in the cyberspace with the consideration of the legal, ethical and regulatory concerns on the battle against hate speech without infringement on free speech. It talks about international and national law systems, court interpretations, and platform accountability systems when it talks about the role of technology in content moderation. The paper concludes by requesting a balance rights-oriented solution to protect the liberty of the individual and digital age peace.*

**Keywords:** *Human Rights, Digital Age, Freedom of Expression, Hate Speech, Cyber Law, Content Moderation.*

---

## 1. INTRODUCTION

The digital age has brought a revolution in communication and provided people a platform like never before to communicate, share their opinion and exchange ideas across the world. The use of social media networks, online forums, and digital platforms has become the most important way



of engaging in the discourse in the society, and it has contributed to democratic participation and sharing of information. Although freedom of expression is a fundamental aspect of democratic societies, its application in the digital realm tends to conflict with the need to reduce the spread of malicious information like hate speech, misinformation, and internet bullying.

Violent discrimination, psychological abuse, and other forms of online hate speech pose a serious danger to human rights, equality, and societal cohesiveness. Politicians, judges, and tech firms have all had to walk a fine line when trying to regulate this kind of content without stifling free speech. Some of the most landmark international human rights documents include the Declaration of Human Rights and the Covenant on Civil and Political Rights. According to the latter, the former ensures the right to free speech, within reasonable limits set for purposes of public order, national security, and the rights of others. This paper will explore the shifting landscape of human rights in the digital age, specifically looking at how laws, case law, and platform policies handle the delicate balance between ensuring free speech online and preventing hate speech.

## **2. HUMAN RIGHTS AND THE DIGITAL AGE**

The human rights that were initially conceived to govern the state-individual relationship are currently experiencing a challenge in cyberspace because the digital spaces are dominated by the private tech firms that manipulate the rights through moderation and data policies. Although the liberty of speech is safeguarded by such instruments as the ICCPR, online freedom of speech remains a complicated issue because of hate speech, disinformation, and different global standards of law, which causes contradictions between free speech and control.

### **2.1. Evolution of Human Rights in Cyberspace**

The history of human rights frameworks was initially created to regulate the interaction between states and individuals in the physical world with issues of freedom, equality and dignity being of concern. Nevertheless, the digital era has changed the situation drastically. In the current world, the role of the quasi-government has been taken over by privately owned technology companies, such as social media giants, search engines, and cloud service providers, which have taken control



of huge digital realms where people communicate, express themselves, and associate. The corporations have never had so much power over the rights of people as a result of content moderation, algorithmic decisions, and data policies.

Certain international documents, like the UDHR and the ICCPR, recognise the right to freedom of speech, privacy, and information as basic human rights. However, these treaties were written before the digital world existed, and they don't specifically include the role of the private sector, so applying them to cyberspace is still a bit of a contentious issue. The United Nations Guiding Principles on Business and Human Rights were put in place to correct this shortcoming; they state that companies should respect human rights, refrain from harming people, and offer solutions in the event that they violate these rights. This has sparked a discussion over whether digital platforms should be considered a public utility or should be allowed to operate autonomously with the freedom to police expression.

## **2.2. Freedom of Expression**

A fundamental component of democratic society, freedom of expression is protected in the constitutions of the majority of democracies and the International Covenant on Civil and Political Rights under Article 19. Also, it ensures that people can freely access any kind of information or ideas, no matter how far away they are. In a well-functioning democracy, the right to free speech includes the ability to criticise the government, express politically charged ideas, and even hold vile or unpopular views.

This privilege, nevertheless, is not without limitations. Legally mandated limitations are permissible under Article 19(3) of the ICCPR if they are essential to safeguard public health, morals, national security, or the reputations or rights of other individuals. In the digital age, this balance becomes much more intricate. The presence of the social media has increased the range and velocity of expression so that such negative content as hate speech, disinformation, and cyber harassment are now rampant. Also, platforms have to cater to a global user base that has varying legal frameworks and cultural standards, which results in uneven content moderation guidelines. As an illustration, what can be regarded as acceptable political speech in one place can be seen as



an illegal hate speech in another. Such international interdependence is posing immense problems of balancing free speech rights with reasonable regulatory interests.

### **3. ONLINE HATE SPEECH: DEFINITION AND CHALLENGES**

When someone uses language, writing, or visuals to incite hostility, violence, or prejudice against another person or group because of their race, religion, ethnicity, gender, sexual orientation, nationality, or any other characteristic, it is considered hate speech. The problem has become even greater in the digital age where online platforms are fast and have a huge outreach. The social media allows negative messages to spread rapidly in a matter of seconds, and the use of pseudonyms or anonymity gives the perpetrators more confidence to spread hate messages, decreasing the sense of accountability and making it difficult to trace the origin of the hate content.

Jurisdictional complexities also complicate the regulation of online hate speech. It is a common phenomenon that the digital content may easily cross national borders, which leads to the conflict of legal standards and cultural norms. Hate speech in one jurisdiction is protected speech in another. Lack of a common international law system leaves a loophole in enforcement, which means that accountability is hard to achieve, and dangerous content can continue to exist on various platforms and in different countries.

The other contentious issue is platform liability. A debate is still raging regarding the role of social media companies as neutral platforms or those that take the responsibility of actively monitoring and removing harmful content. When it comes to user-generated content, platforms in the United States are heavily regulated under Section 230 of the Communications Decency Act, whereas in the European Union, the Digital Services Act imposes stricter moderation standards. Politicians must tread carefully to avoid censorship and the suppression of authentic debates by striking a balance between the rights to free speech and the control of hate speech.



#### **4. LEGAL FRAMEWORK AND JUDICIAL INTERPRETATIONS**

The international hate speech regulation takes into account the balance between free expression and the prevention of harm as with the frameworks of the ICCPR and the ECHR and tools such as the six-part test of the UN Rabat Plan. Whereas the U.S. safeguards the majority of hate speech under the First Amendment unless it leads to an imminent violence, EU has stringent removal conditions on online hate speech. In India, Article 19(1) (a) protects speech, but Article 19(2) allows reasonable restrictions, and the hate speech on the internet is regulated by such legislation as IPC Section 153A and 295A and IT Rules, 2021, which exemplify the varying global attitudes towards protecting democratic principles and maintaining order in the community.

##### **4.1. International Perspective**

Finding a middle ground between free expression and non-harm is the goal of international hate speech regulation. Article 19 of the ICCPR establishes the right to free speech, while Article 20 prohibits inciting hate based on nationality, race, or religion in a way that discriminates, is hostile, or violent (2). Thus, it is a timely reminder that safeguarding pluralistic societies from harmful narratives should not come at the expense of democratic freedoms.

In a similar vein, the right to free expression is guaranteed in Article 10 of the European Convention on Human Rights, with the exception of certain limitations that are fundamental to preserving democratic society's public order and security. The European Court of Human Rights often approves restrictions on hate speech when it threatens democratic ideals or physical violence.

A six-factor threshold test was recommended by the UN Rabat Plan of Action (2012) to help with implementation and this includes the consideration of context, status of the speaker, intent, content, reach, and probability of harm. This will be one of the ways of avoiding the utilization of hate speech laws to censor speech and the culpability of those who use damaging and inciting speech.



#### 4.2. National Frameworks

National approaches to regulating hate speech reflect diverse constitutional traditions and legal philosophies:

- **United States:** The First Amendment secures strong safeguards of freedom of speech, such as offensive and hateful speech, unless it directly provokes imminent lawless activity (as determined in the case of *Brandenburg v. Ohio*). This is a high bar so that most types of hate speech are legal unless they are explicitly related to violence and America is thus an outlier in the European context.
- **European Union:** Big platforms are obligated to filter hate speech items within 24 hours of receiving a notification, according to the EU Code of Conduct on Countering Illegal Hate Speech Online, which is enforced by the EU with greater vigour. They can be subject to some penalties in case of refusal to do it because a more serious focus on combating online hatred is envisaged.
- **India:** Although freedom of speech and expression is guaranteed in Article 19(1)(a) of the Indian Constitution, reasonable restrictions can be imposed in Article 19(2) to protect public order, decency, morality, and security. Finding a middle ground between these ideas has been an aim of Indian jurisprudence. The Supreme Court maintained limitations meant to preserve public order in the well-known case of *Shreya Singhal v. Union of India* (2015), but found Section 66A of the IT Act to be arbitrary in its censoring because it was too wide and imprecise.

The Indian Penal Code (Sections 153A, 295A) makes it a crime to conduct a religious offence or to express hatred. The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, which have been characterised as either legislation or censorship, required platforms to delete illegal content within very short periods of time.



## **5. FREEDOM OF EXPRESSION VS REGULATION OF HATE SPEECH**

Digital age has introduced a basic paradox, the preservation of the freedom of speech that is a pillar of democracy and the curtailment of hate speech that undermines equality and social cohesion. Excessive regulation threatens to suppress legitimate speech and allow the authoritarian abuse of the power, and insufficient regulation contributes to discrimination, violence, and radicalization, endangering vulnerable groups and democratic confidence. Courts are having difficulty establishing definite limits to unlawful incitement because the conventional tests cannot reflect the pace, proportions, and anonymity of the internet, so context-intent, reach, and likelihood of harm is paramount but difficult to measure across jurisdictions.

The use of AI-driven content moderation complicates the situation, creating the problems of bias, accuracy, and transparency. Automated systems are prone to misinterpreting satire, disproportionately silencing marginalized voices, and without due process, which has led to demands of algorithmic accountability and human oversight. These conflicting interests need to be balanced with a multi-stakeholder approach which involves legal protection, human rights values and technological controls to provide a strong protection of speech freedom and protection of vulnerable groups to harmful speech.

## **6. CONCLUSION**

With the rise of the digital age came new opportunities and risks to free speech, completely altering the nature of human rights. While the internet does great things for fostering democratic engagement and worldwide discourse, it also has the potential to give rise to hate speech, which threatens societal harmony, equality, and human decency. One of the most important challenges that governments, courts, and digital corporations throughout the world are confronting is determining how to ban harmful content while still protecting free speech. While the promotion of violence and discrimination will persist in the absence of speech regulation, the risk of overly controlling speech is that it will stifle free expression. A complex, rights-based strategy grounded on clearly defined legal concepts, proportionality, judicial review, and transparent platform governance is necessary to uphold democratic principles and human rights in the digital era.



## REFERENCES

1. Alkiviadou, N. (2019). *Hate speech on social media networks: towards a regulatory framework?*. *Information & Communications Technology Law*, 28(1), 19-35.
2. Castaño-Pulgarín, S. A., Suárez-Betancur, N., Vega, L. M. T., & López, H. M. H. (2021). *Internet, social media and online hate speech. Systematic review. Aggression and violent behavior*, 58, 101608.
3. Chetty, N., & Alathur, S. (2018). *Hate speech review in the context of online social networks. Aggression and violent behavior*, 40, 108-118.
4. Howard, J. W. (2019). *Free speech and hate speech. Annual Review of Political Science*, 22(1), 93-109.
5. MacAvaney, S., Yao, H. R., Yang, E., Russell, K., Goharian, N., & Frieder, O. (2019). *Hate speech detection: Challenges and solutions. PloS one*, 14(8), e0221152.
6. Narotra, R. (2024). *Hate Speech vs. Freedom of Expression: Striking the Right Balance in a Diverse Society. LawFoyer Int'l J. Doctrinal Legal Rsch.*, 2, 341.
7. Paz, M. A., Montero-Díaz, J., & Moreno-Delgado, A. (2020). *Hate speech: A systematized review. Sage Open*, 10(4), 2158244020973022.
8. Strossen, N. (2018). *Hate: Why we should resist it with free speech, not censorship. Oxford University Press*.
9. Thuku, J. T., & Mbaaro, M. (2022). *Hate speech legislation and freedom of expression: Finding the balance. Interdisciplinary Studies in Society, Law, and Politics*, 1(2), 26-33.
10. Tontodimamma, A., Nissi, E., Sarra, A., & Fontanella, L. (2021). *Thirty years of research into hate speech: topics of interest and their evolution. Scientometrics*, 126(1), 157-179.



## Author's Declaration

I as an author of the above research paper/article, here by, declare that the content of this paper is prepared by me and if any person having copyright issue or patent or anything otherwise related to the content, I shall always be legally responsible for any issue. For the reason of invisibility of my research paper on the website /amendments /updates, I have resubmitted my paper for publication on the same date. If any data or information given by me is not correct, I shall always be legally responsible. With my whole responsibility legally and formally have intimated the publisher (Publisher) that my paper has been checked by my guide (if any) or expert to make it sure that paper is technically right and there is no unaccepted plagiarism and hentriacontane is genuinely mine. If any issue arises related to Plagiarism/ Guide Name/ Educational Qualification /Designation /Address of my university/ college/institution/ Structure or Formatting/ Resubmission /Submission /Copyright /Patent /Submission for any higher degree or Job/Primary Data/Secondary Data Issues. I will be solely/entirely responsible for any legal issues. I have been informed that the most of the data from the website is invisible or shuffled or vanished from the database due to some technical fault or hacking and therefore the process of resubmission is there for the scholars/students who finds trouble in getting their paper on the website. At the time of resubmission of my paper I take all the legal and formal responsibilities, If I hide or do not submit the copy of my original documents (Andhra/Driving License/Any Identity Proof and Photo) in spite of demand from the publisher then my paper maybe rejected or removed from the website anytime and may not be consider for verification. I accept the fact that as the content of this paper and the resubmission legal responsibilities and reasons are only mine then the Publisher (Airo International Journal/Airo National Research Journal) is never responsible. I also declare that if publisher finds Any complication or error or anything hidden or implemented otherwise, my paper maybe removed from the website or the watermark of remark/actuality maybe mentioned on my paper. Even if anything is found illegal publisher may also take legal action against me

**Pallavi Chaurasia**

\*\*\*\*\*