

# INTERNATIONAL JOURNAL FOR LEGAL RESEARCH AND ANALYSIS



Open Access, Refereed Journal Multi-Disciplinary  
Peer Reviewed

[www.ijlra.com](http://www.ijlra.com)

## **DISCLAIMER**

No part of this publication may be reproduced or copied in any form by any means without prior written permission of Managing Editor of IJLRA. The views expressed in this publication are purely personal opinions of the authors and do not reflect the views of the Editorial Team of IJLRA.

Though every effort has been made to ensure that the information in Volume II Issue 7 is accurate and appropriately cited/referenced, neither the Editorial Board nor IJLRA shall be held liable or responsible in any manner what sever for any consequences for any action taken by anyone on the basis of information in the Journal.

Copyright © International Journal for Legal Research & Analysis

## EDITORIALTEAM

### EDITORS

#### **Dr. Samrat Datta**

*Dr. Samrat Datta Seedling School of Law and Governance, Jaipur National University, Jaipur. Dr. Samrat Datta is currently associated with Seedling School of Law and Governance, Jaipur National University, Jaipur. Dr. Datta has completed his graduation i.e., B.A.LL.B. from Law College Dehradun, Hemvati Nandan Bahuguna Garhwal University, Srinagar, Uttarakhand. He is an alumnus of KIIT University, Bhubaneswar where he pursued his post-graduation (LL.M.) in Criminal Law and subsequently completed his Ph.D. in Police Law and Information Technology from the Pacific Academy of Higher Education and Research University, Udaipur in 2020. His area of interest and research is Criminal and Police Law. Dr. Datta has a teaching experience of 7 years in various law schools across North India and has held administrative positions like Academic Coordinator, Centre Superintendent for Examinations, Deputy Controller of Examinations, Member of the Proctorial Board*



#### **Dr. Namita Jain**



*Head & Associate Professor*

*School of Law, JECRC University, Jaipur Ph.D. (Commercial Law) LL.M., UGC-NET Post Graduation Diploma in Taxation law and Practice, Bachelor of Commerce.*

*Teaching Experience: 12 years, AWARDS AND RECOGNITION of Dr. Namita Jain are - ICF Global Excellence Award 2020 in the category of educationalist by I Can Foundation, India. India Women Empowerment Award in the category of "Emerging Excellence in Academics by Prime Time & Utkrish Bharat Foundation, New Delhi. (2020). Conferred in FL Book of Top 21 Record Holders in the category of education by Fashion Lifestyle Magazine, New Delhi. (2020). Certificate of Appreciation for organizing and managing the Professional Development Training Program on IPR in Collaboration with Trade Innovations Services, Jaipur on March 14th, 2019*

## Mrs.S.Kalpna

Assistant professor of Law

*Mrs.S.Kalpna, presently Assistant professor of Law, VelTech Rangarajan Dr.Sagunthala R & D Institute of Science and Technology, Avadi. Formerly Assistant professor of Law,Vels University in the year 2019 to 2020, Worked as Guest Faculty, Chennai Dr.Ambedkar Law College, Pudupakkam. Published one book. Published 8Articles in various reputed Law Journals. Conducted 1Moot court competition and participated in nearly 80 National and International seminars and webinars conducted on various subjects of Law. Did ML in Criminal Law and Criminal Justice Administration.10 paper presentations in various National and International seminars. Attended more than 10 FDP programs. Ph.D. in Law pursuing.*



## Avinash Kumar



*Avinash Kumar has completed his Ph.D. in International Investment Law from the Dept. of Law & Governance, Central University of South Bihar. His research work is on "International Investment Agreement and State's right to regulate Foreign Investment." He qualified UGC-NET and has been selected for the prestigious ICSSR Doctoral Fellowship. He is an alumnus of the Faculty of Law, University of Delhi. Formerly he has been elected as Students Union President of Law Centre-1, University of Delhi. Moreover, he completed his LL.M. from the University of Delhi (2014-16), dissertation on "Cross-border Merger & Acquisition"; LL.B. from the University of Delhi (2011-14), and B.A. (Hons.) from Maharaja Agrasen College, University of Delhi. He has also obtained P.G. Diploma in IPR from the Indian Society of International Law, New Delhi. He has qualified UGC – NET examination and has been awarded ICSSR – Doctoral Fellowship. He has published six-plus articles and presented 9 plus papers in national and international seminars/conferences. He participated in several workshops on research methodology and teaching and learning.*

## **ABOUT US**

INTERNATIONAL JOURNAL FOR LEGAL RESEARCH & ANALYSIS ISSN- 2582-6433 is an Online Journal is Monthly, Peer Review, Academic Journal, Published online, that seeks to provide an interactive platform for the publication of Short Articles, Long Articles, Book Review, Case Comments, Research Papers, Essay in the field of Law & Multidisciplinary issue. Our aim is to upgrade the level of interaction and discourse about contemporary issues of law. We are eager to become a highly cited academic publication, through quality contributions from students, academics, professionals from the industry, the bar and the bench. INTERNATIONAL JOURNAL FOR LEGAL RESEARCH & ANALYSIS ISSN 2582-6433 welcomes contributions from all legal branches, as long as the work is original, unpublished and is in consonance with the submission guidelines.

# **BALANCING RIGHT TO FREE SPEECH AND RIGHT TO REPUTATION IN THE DIGITAL ERA: A LEGAL ANALYSIS OF AI AND CYBER DEFAMATION LAWS**

AUTHORED BY - PRIA MAKANDA\* & PROF. DR. CHANDRAKANTHI. L\*\*

## **ABSTRACT**

Humans have evolved as social animals far beyond history. While freedom to speak and express one's opinions were considered natural rights in prehistoric times, they were only acknowledged as fundamental human rights globally soon after several revolutionary moments by freedom fighters throughout history. We are all aware that free speech is accompanied by the absolute freedom to express one's thoughts; nevertheless, it further imposes constraints on what can and cannot be said, which would jeopardise the reputation of an individual. The right to reputation has been widely referred to as an inherent human right since ancient times. The role of the law is to strike a balance between the freedom of speech and expression while not breaching the right to reputation.

The advent and adoption of technological advancements in Internet Cyberspace have increased tremendously over the years. Social networking sites and online communities have created an extensive platform enabling anybody to share anything from every corner of the globe. Nevertheless, the virtual world has frequently failed in safeguarding an individual's reputation online, too. The recent development of artificial intelligence (AI) has proved to be an opportunity and a disadvantage for communication over the internet, posing emerging problems in harmonising free speech and reputation.

The ethical and legal implications of cyber defamation in an AI-driven society are examined in this research paper. Firstly, the authors begin by discussing the legal facets of India's cyber defamation laws from both a constitutional and international standpoint. Second, discussing how AI may improve free expression while simultaneously posing dangerous hazards to

---

\* Researcher Scholar, University Law College & Dept of Studies in Law Jnana Bharathi Campus, Bangalore University Bengaluru – 560 056

\*\* Professor , Faculty of Law , University Law College & Dept of Studies in Law Jnana Bharathi Campus, Bangalore University Bengaluru – 560 056

people's online reputations. Third, the judiciary's function is to achieve equilibrium between the two rights. Finally, to lessen the possible harm that defamation in the age of the internet might do to society, the authors of this article desire to add to the current deliberation on the crucial importance of a legal framework and proper implementation of AI (artificial intelligence).

**Keywords:** Cyber Defamation, Artificial Intelligence, Reputation, Free Speech

## 1. Introduction:-

A person's reputation is not merely a superficial attribute; rather, it is an integral part of their character and integrity, which is essential to living a life of honour and respect in society. It is, however, a delicate entity, susceptible to external influence and beyond our direct control. This emphasises how crucial it is to protect one's reputation because it has a significant influence on one's status and place in society. It underlines the necessity to safeguard one's reputation, as it can gravely impact one's standing and position in society.

The 21<sup>st</sup> century is one where the nature and pattern of communication have changed completely. It is possible to share content with a wide audience, spread information quickly, and reach a mass population. But as it is said, there are two sides to a coin. It is possible to disseminate misinformation about someone with the intent to cause serious harm to their reputation through libelous content on online platforms. Many do not hesitate to post and share content that can cause serious harm to someone. This might be done intentionally by someone, but in an age where people mindlessly scroll and share, rumours do not take very long to spread. Online spaces have become safe havens for defamatory claims. In such a situation, it is essential to know and understand the ways to protect your reputation online. Although India's legal system provides various avenues for those who have been defamed to seek a remedy, the extent and immediacy of digital communication make it difficult to enforce laws and hold people accountable for such acts. To reduce the dangers of online defamation, it is crucial to comprehend the intricacies of defamation law and take preventive action to protect one's online reputation.

The law on defamation was never naïve to human history. All the current laws have been codified by Britain & the United States, but their traces can be rooted in the Roman, Greek, Indian, and Indian civilizations, encompassing several punishments for any form of insults or

damage caused by defamation.

The common law action grew out of the English ecclesiastical courts' failure to deal satisfactorily with defamation. While the church courts could order offenders to apologize, victims often found such remedies inadequate and turned to duels for satisfaction. The Scandalum Magnatum, passed in 1275 to stop this violence, introduced two justifications for defamation law that remain relevant.<sup>1</sup>

By making some public statements unlawful, however, defamation law runs counter to another widely accepted legal tenet- the right to freedom of expression. Constitutions from various countries drafted from the eighteenth century to the present contain provisions that guarantee free speech within states. On the international level, treaties, such as the International Covenant on Civil and Political Rights and the European Convention on Human Rights, protect the right to free expression. Article 19 of the Universal Declaration of Human Rights states, "Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers

Defamation has been defined as the 'intentional infringement of another's right to his good name, or, more comprehensively, the wrongful, intentional publication of words or behaviour concerning another which has the tendency to undermine his status, good name.

Professor Prosser defines defamation as: "That which tends to injure 'reputation' in the popular sense; to diminish the esteem, respect, goodwill, or confidence in which plaintiff is held, or to excite adverse, derogatory or unpleasant feelings against him"<sup>2</sup>

Similarly, any such act taking place in cyberspace leads to cyber defamation or online defamation. Cyber defamation occurs when a computer connected to the internet is used as a tool or a medium to defame a person or an entity.

---

<sup>1</sup>Ashu M. G. Solo, *Combating Online Defamation and Doxing in the United States*, Proceedings of the 2019 International Conference on Internet Computing and Internet of Things, Las Vegas, July 2019 [https://www.researchgate.net/publication/334604707\\_Combating\\_Online\\_Defamation\\_and\\_Doxing\\_in\\_the\\_United\\_States/references](https://www.researchgate.net/publication/334604707_Combating_Online_Defamation_and_Doxing_in_the_United_States/references) Last visited on 10/12/2024 .

<sup>2</sup> Hurwitz, Gus, "Defamation and Privacy" (2024). *Book Chapters*. 458. [https://scholarship.law.upenn.edu/faculty\\_chapters/458](https://scholarship.law.upenn.edu/faculty_chapters/458) <https://doi.org/10.1017/9781009174411.009>.

## 2. EVOLUTION OF CYBER DEFAMATION

Defamation has existed for a long time; the only difference is that, with the changing times, the ways of doing things have changed a lot. With the development of technology, various new ways of defamation have evolved. Cyber defamation evolved right from the evolution of the cyber age in the early 1990s. From then on, various laws in almost all countries have been made considering the widening scope of cyber defamation.

Cyber Defamation is of 2 types, viz, firstly, the individual who primarily publishes defamatory statements on the internet, such as email writers and website content providers, is liable for online defamation. These individuals are the primary publishers of defamatory content that is produced and published in cyberspace for Cyber Defamation. Secondly, there is an increasing number of websites that view various types of information, which are established and controlled by ISPs (Internet Service Providers). They are responsible for any defamatory statement published on their site because they must review the material before it is made public. Essentials to prove Cyber Defamation. There are three requirements that must be fulfilled to prove Defamation. They are-

- The imputation must be published
- Imputation must be in the form of noticeable representation.
- Making such imputations should be done with the intent of causing damage or with the understanding that it would harm the person's goodwill.

In India, Section 499 of the Indian Penal Code primarily governs the law on defamation; however, it is pertinent to note that the law has been extended to “electronic documents”. Section 469 of the IPC (forgery for purpose of harming reputation) has been amended by the Information Technology Act, 2000 to include ‘electronic record forged’ and now reads as a whole as – whoever commits forgery, intending that the document or electronic record forged shall harm the reputation of any party, or knowing that it is likely to be used for that purpose, shall be punished with imprisonment of either description for a term which may extend to three years, and shall also be liable to fine.<sup>3</sup>

Our growing reliance on the Internet and social networking sites has created various legal challenges in the country. When it comes to defamation, one of the biggest issues is identifying

---

<sup>3</sup> **Rishita Naredi**, *A Critical Analysis of Cyber Defamation Laws in India*, 8 INT'L J. INNOVATIVE RES. TECH. 301 (2021). [https://ijirt.org/publishedpaper/IJIRT151521\\_PAPER.pdf](https://ijirt.org/publishedpaper/IJIRT151521_PAPER.pdf)

the person responsible for harming our reputation or figuring out who has seen the defamatory statement. This is especially difficult with web pages like blogs or other media outlets such as newspapers or magazines. Bloggers may choose to remain anonymous or use pseudonyms to protect their identities.

This makes it difficult to determine who published the statement if it appears on someone's blog. It becomes even harder to trace readers who leave comments on blogs or online news articles, as most sites don't require users to provide their real names or any personal information like email or location. Even when sites do request this information, people can use false identities. This makes it difficult to track these individuals. Additionally, when defamatory statements are published on platforms like Facebook, they can spread rapidly and reach a large audience, causing significant damage to the person being targeted.

### **3. VARIOUS LEGAL PROVISIONS DEALING WITH DEFAMATION & CYBER DEFAMATION IN INDIA**

There are several provisions for online defamation, and they vary depending on the type of offense committed by the offender. Cyber Defamation has its liabilities in many laws. These laws are stated below.

#### **3.1 INDIAN PENAL CODE, 1860**

- i) Section 499 According to Section 499 of IPC,1860, "whoever, by words either spoken or read, or by signs or by any visible representation, makes or publishes any imputation concerning any person intending to harm, or knowing or having reason to believe that such imputation will harm, the reputation of such person, is said, to defame that person." The law, after the introduction of the IT Act, 2000, included "Speech" and "Documents" in electronic form. Along with Section 500, this section in IPC deals with the punishment of the above-mentioned law of defamation. As per Section 500, "Whoever defames another is punishable with imprisonment for two years, a fine, or both."
- ii) Section 503 This section deals with offenses committed through the use of computer devices, such as sending emails, posting messages, and leaving comments, to harm someone's reputation or threaten them.

- iii) Section 506 The penalty for criminal intimidation is discussed in this section. As per Section 506, whoever commits the above-mentioned offense shall be punished with 2-year' jail term, a fine, or both<sup>4</sup>.

**Defamation as per the Bharatiya Nyaya Sanhita (BNS) 2023:-** Defamation is making or publishing false and harmful statements about an individual or entity that damage their reputation. Defamation can occur in libel (written defamation) or slander (spoken defamation).

• (Criminal & Civil):  
o Criminal Defamation: Under Section 356 (1) of the Bharatiya Nyaya Sanhita, 2023, defamation is a criminal offence, and Section 356(2) prescribes punishment for it, which can include imprisonment for up to two years, a fine, or both. Criminal defamation involves the communication of a false statement that damages someone's reputation. Civil Defamation: Defamation is also an actionable civil wrong (tort) in India. A person whose reputation has been damaged by a defamatory statement can file a lawsuit seeking compensation for the harm caused. Defamation law in India allows for certain exceptions, including statements made in good faith, truth spoken for the public good, fair criticism of public performance, and judicial proceedings. These fall under "reasonable restrictions" permitted under Article 19(2) of the Indian Constitution, which allows limits on the freedom of speech<sup>5</sup>.

### 3.2. Information Technology Act, 2000

- i) Section 66 Although Section 66A of the Information Act of 2000 does not explicitly address cyber defamation, it does make the act of sending defamatory material for causing insult, injury, or criminal intimidation punishable.
- ii) Section 79 States that an intermediary is not liable if they do not initiate or modify defamatory content, but merely act as a facilitator.

After the case of **Shreya Singhal and Ors, vs Union of India**, this section of the IT Act, 2000, was struck down as it was considered to violate the Fundamental Rights of the Indian Constitution<sup>6</sup>.

<sup>4</sup> **Samridhi M**, *IPC Notes: Defamation Under the Indian Penal Code*, LACTOPUS, <https://lawctopus.com/clatalogue/clat-ug/ipc-notes-defamation-under-the-indian-penal-code/> (last visited Dec. 13, 2024).

<sup>5</sup> *Harm to Reputation: Introduction to Defamation Under BNS*, LAWGICAL SHOTS, <https://lawgicalshots.com/harm-to-reputation-introduction-to-defamation-under-bns/> (last visited Dec. 13, 2024).

<sup>6</sup> **Shivanshu**, *Cyber Defamation: What Is, Laws and Protection*, INTELLIPAAT, <https://intellipaata.com/blog/what-is-cyber-defamation/> (last visited Dec. 15, 2024).

## 4. CYBER DEFAMATION IN THE UNITED STATES

The United States protects free speech under the **First Amendment**, making defamation cases more challenging. However, defamation laws in the U.S. vary by state, and victims of cyber defamation can still seek recourse.

- **The Communications Decency Act (CDA), Section 230:** This is a crucial law in U.S. internet regulation, as it provides immunity to online platforms from being held liable for defamatory content posted by third parties. However, individuals who post defamatory statements can still be held responsible.
- **Public Figures vs. Private Individuals:** In the U.S., there is a distinction between public figures and private individuals when it comes to defamation. Public figures must prove "actual malice" (i.e., the statement was made with knowledge of its falsity or with reckless disregard for the truth), whereas private individuals only need to prove negligence.

### 4.2 CYBER DEFAMATION IN THE UNITED KINGDOM

The UK has stringent defamation laws under the **Defamation Act 2013**, which applies to both offline and online defamation. The key provisions include:

- **Serious Harm Threshold:** To bring a defamation claim, the claimant must prove that the defamatory statement caused, or is likely to cause, serious harm to their reputation. In the case of businesses, serious financial loss must be demonstrated.
- **Responsibility of Website Operators:** The Defamation Act 2013 provides a defense for website operators if they can prove they didn't post the defamatory content themselves. However, they are required to act upon complaints and remove defamatory content once notified<sup>7</sup>.

## 5. BALANCING FREE EXPRESSION AND REPUTATION IN THE INDIAN CONTEXT

In India, the judiciary plays a crucial role in balancing the right to free expression, enshrined under Article 19(1)(a) of the Constitution, and the right to reputation, which is part of the Right to Life under Article 21. Both rights are fundamental, but neither is absolute, and the judiciary has developed a nuanced framework to ensure equilibrium between them.

---

<sup>7</sup> Understanding Cyber Defamation: Types, Legal Frameworks, and Remedies For Victims. <https://restthecase.com/knowledge-bank/cyber-defamation> last visited on 11/12/2024

- a. Article 19(1)(a): Freedom of Speech and Expression:- The Constitution guarantees citizens the right to express their opinions freely. This includes the right to dissent, criticize the government, and participate in public discourse. However, Article 19(2) imposes reasonable restrictions, including those for maintaining the dignity and reputation of individuals.
- b. Article 21: Right to Life and Reputation:- The Supreme Court of India has interpreted the Right to Life expansively to include the right to live with dignity, which encompasses the protection of one's reputation. In cases where reputational harm occurs due to the misuse of free speech, the courts step in to balance the two rights.

Under Indian law, reputation and defamation are distinct concepts. Reputation is a protected right, and defamation is an actionable offence.

Reputation: Reputation refers to the esteem or standing of an individual in the eyes of society. It encompasses the public perception of a person's character, integrity, and overall image. Reputation is implicitly protected under Article 21 of the Indian Constitution, which guarantees the right to life and personal liberty. The Supreme Court of India has consistently held that reputation is a vital part of a person's dignity and is protected as part of their right to life. While reputation itself is not actionable unless harmed, individuals can seek legal remedies when their reputation is damaged through defamatory actions. This protection is enforced primarily through defamation laws.

The right to freedom of speech and expression, as enshrined in Article 19(1)(a) of the Indian Constitution, is a cornerstone of democratic governance. It allows individuals to freely express their thoughts, dissent, and opinions, which are essential for personal growth, societal development, and a healthy democracy.<sup>8</sup>

The Supreme Court, while deciding the case of **Subramaniam Swamy v Union of India**, observed in its judgment that the court in this case heavily stressed the role of reputation within the right to life as enshrined under Article 21, recognising it as a fundamental right. The necessity to balance the fundamental rights is paramount, as the court cautioned against sacrificing one's reputation on the altar of another's freedom of speech. The court examined defamation not only in the light of a civil wrong but also as a criminal offence by elucidating

---

<sup>8</sup> Saurabh Mishra, *Balancing Act: Free Speech v Online Content Moderation*, 3 INTEGRAL L. REV. 1 (2024–2025). [https://www.iul.ac.in/LawJournal/Pdf/Volume\\_III\\_Manuscript\\_10.pdf](https://www.iul.ac.in/LawJournal/Pdf/Volume_III_Manuscript_10.pdf)

that, “causing harm to the reputation of a person is the basis on which the offence is founded and mens rea is a condition precedent to constitute the same offence”.

While AI technologies have the potential to completely transform and revolutionise various aspects of our lives, like healthcare, transportation, and communication, their rapid development also brings up serious issues with fairness, privacy, and autonomy, one of the primary challenges comes from the inherent biases in AI algorithm, which reinforces inequality and discrimination, especially when it comes to the decision making process in the criminal justice system. The increasing reliance on AI in every field of human development has raised concerns about privacy and government overreach. The wide accessibility of such technology could facilitate hateful, sexual, and violent imagery and videos targeting both public and private individuals alike. AI technologies have therefore been wielded as potent tools capable of inflicting significant harm, from creating lifelike videos, images, and voice mimicry. AI manipulation has transcended traditional boundaries, posing profound threats to personal integrity and societal trust<sup>9</sup>.

### **5.1 THE JUDICIARY’S ROLE IN BALANCING THE RIGHTS**

Various cases involved the issue that defamation law unfairly imposes restrictions on the freedom of speech, such as.

In the case of *Subramaniam Swamy vs. the Union of India*, the petitioner claimed that trying to limit the sharing of public views, thoughts, and critiques through criminal legal actions would harm the functioning of the democratic system. Nonetheless, the Supreme Court maintained that the legal right to free speech does not permit an individual to tarnish the reputation of another by defaming them.

In another case of *Rajagopal v. State of Tamil Nadu*. The issue of whether civil defamation laws were constitutional and their compatibility with the right to free speech was explored. The court determined that the concept of common law defamation is limited under Article 19(1)(a) due to the undue benefit it provided by shifting the burden of proof to the fault. The primary criticism of Section 499 was that it made a private matter a criminal offense. This law thus imposed a restriction on the freedom of expression. To hold an individual responsible for civil

---

<sup>9</sup> Saksham Pradyot, Interplay of Free Speech, Defamation, and Reputation: Constitutional Principles and State Duties in Legal Interpretation in India, 4 J. INFORMATICS EDUC. & RES. 3 (2024).

defamation, it is necessary to demonstrate that the accused acted with a complete disregard for the truth in committing defamation.

In the most landmark *Shreya Singhal v. Union of India*, the Supreme Court of India has attempted to strike a balance between the rights protected in Article 19(1)(a) and the justifiable limitations allowed under Article 19(2). In its judgment remarks, the Supreme Court has accurately and distinctly stated that “in democracy, the freedom of thought and expression is a fundamental principle that holds great importance within our constitutional framework.”

Moreover, the Court has made it clear that there is a difference between discussions, advocacy, and incitement. The court has determined that justifiable limitations on free speech and expression under Article 19(2) can only be applied when there is clear evidence of incitement.

## 6. JUDICIAL RESPONSE TO THE OFFENCE OF CYBER DEFAMATION

In the first case of cyber defamation in India, *SMC Pneumatics (India) Pvt. Ltd. v. Jogesh Kwatra*, the reputation of a corporate was being defamed by an employee of the plaintiff company by sending derogatory, defamatory, obscene, emails obscene, vulgar, filthy and abusive emails to its employers and also to different subsidiaries of the said company all over the world with the aim to defame the company and its Managing Director. The Hon'ble Judge of the Delhi High Court passed an ex parte ad interim injunction observing that a prima facie case had been made out by the plaintiff<sup>10</sup>.

Consequently, the Delhi High Court restrained the defendant from sending derogatory, defamatory, obscene, vulgar, humiliating, and abusive emails either to the plaintiffs or to its sister subsidiaries all over the world, including their Managing Directors and their Sales and Marketing departments. Further, the Hon'ble Judge also restrained the defendant from publishing, transmitting, or causing to be published any information in the actual world, as well as in cyberspace, which is derogatory, defamatory, or abusive of the plaintiffs.

In another case, *State of Tamil Nadu v. Suhas Katti*, related to the posting of obscene, defamatory, and annoying messages about a divorced woman in the Yahoo message group. E-

---

<sup>10</sup> **R. Rajagopal v. State of Tamil Nadu**, (1994) 6 S.C.C. 632 (India).  
**Shreya Singhal v. Union of India**, (2015) 5 S.C.C. 1 (India).

mails were also forwarded to the victim for information by the accused through a false e-mail account opened by him in the name of the victim. The posting of the message resulted in annoying phone calls to the lady, in the belief that she was soliciting. Based on a complaint made by the victim in February 2004, the Police traced the accused to Mumbai and arrested him within the next few days. Relying on the expert witnesses and other evidence produced before it, including the witnesses of the Cyber Cafe owners, the Additional Chief Metropolitan Magistrate held the accused guilty of offences under section 469, 509 IPC and 67 of IT Act, 2000 and the accused is convicted and is sentenced for the offence to undergo RI for 2 years under 469 IPC and to pay fine of Rs.500/-and for the offence u/s 509 IPC sentenced to undergo 1 year Simple imprisonment and to pay fine of Rs.500/- and for the offence u/s 67 of IT Act 2000 to undergo RI for 2 years and to pay fine of Rs.4000/- All sentences to run concurrently.” The conviction of the accused was achieved successfully within a relatively quick time of 7 months from the filing of the FIR.

In the case of *State of Tamil Nadu v. Suhas Katti*, the Delhi High Court has held that publication is a comprehensive term, embracing all forms and media, including the Internet. That an internet publication has wider viewership, or a degree of permanence, and greater accessibility, than other fixed (as opposed to intangible) media of expression does not alter the essential part, i.e., that it is a forum or medium. There is much sense in having more defined criteria, taking into account the nature of the internet content. Injunctions on internet content should not be readily granted (especially ex parte) since, firstly, the internet is an easy, self-publishing platform providing a medium of expression for marginal individuals not having corporatist outlets. Secondly, the internet facilitates the distribution of content for a minor cost to a vast audience. Both the alleged injury and the concern about free speech are greater due to the wider dissemination of the content. These are only some of the concerns that set the internet apart, and it is desirable to have a nuanced appreciation<sup>11</sup>.

## 7. AI'S ROLE IN FREE EXPRESSION AND ONLINE REPUTATION

Artificial Intelligence (AI) is reshaping how people express themselves online and how reputations are built, maintained, or tarnished. While AI opens up unprecedented opportunities for creative freedom and democratic discourse, it also introduces serious risks to individual and

---

<sup>11</sup> □ **SMC Pneumatics (India) Pvt. Ltd. v. Jogesh Kwatra**, (2004) (India) (Del. HC) (Unreported case but widely cited for cyber defamation principles).

□ **State of Tamil Nadu v. Suhas Katti**, (2004) (India) (Chennai C.J.M. Ct.) (Landmark case for conviction in cybercrime).

collective reputations, which are mentioned below.

- 1. Accessible and Scalable Content Creation:-** Generative AI tools, such as ChatGPT, DALL·E, and similar systems, allow users to create text, images, music, and videos effortlessly. These tools democratize content creation, giving individuals without professional skills or resources the ability to express their ideas artistically or academically. This has made creative expression and public discourse more inclusive.
- 2. Language Translation and Bridging Global Divides:-** AI-driven translation systems like Google Translate and DeepL have enabled people from diverse linguistic backgrounds to communicate and share ideas. For example, an activist in a rural area speaking a local language can now share their message with a global audience without relying on intermediaries. This improves cross-cultural understanding and fosters a global exchange of ideas.
- 3. Empowering Marginalized Voices:-** AI can identify and promote underrepresented content creators by analyzing patterns and recommending diverse voices to users. In environments where certain opinions are suppressed, AI-powered anonymity tools also allow individuals to share views without fear of retaliation, contributing to a more balanced and inclusive conversation.
- 4. Personalized Platforms for Expression:-** Social media platforms use AI to create personalized feeds, helping users connect with communities that share their interests. By filtering out irrelevant content, AI enables individuals to engage in meaningful conversations, contributing to niche movements or fostering collective action on important issues.
- 5. AI-Driven Moderation for Safer Expression:-** AI moderation tools play a critical role in maintaining safe online environments, identifying and removing harmful content like hate speech, threats, or harassment. By doing so, these systems ensure that more voices can participate without fear of intimidation or abuse.

### 7.1 The Hazards AI Poses to Online Reputations

While AI enhances free expression, it also poses significant threats to online reputations, often amplifying the speed and scale at which reputational harm can occur.

- 1. Deepfake Technology:-** One of the most alarming risks posed by AI is the creation of deepfakes—realistic yet entirely fabricated videos, images, or audio. Deepfakes can be weaponized to create misleading or defamatory content, such as a fake video of a public figure engaging in unethical behavior. The damage caused by such false content can be

- irreparable, especially if it goes viral before being debunked.
2. **Misinformation and Manipulation:-** AI can generate realistic but false information at scale. This includes fake news articles, doctored images, and other forms of disinformation that can ruin reputations, influence elections, or harm public trust. Tools like generative AI make it easier for bad actors to spread damaging narratives about individuals or organizations.
  3. **Algorithmic Bias and Amplification:-** AI algorithms, which decide what content to display or suppress, may inadvertently harm reputations by amplifying negative or misleading content. For example, a minor controversy about an individual can be exaggerated by algorithms prioritizing sensational content, creating a disproportionate backlash.
  4. **Identity Theft and Impersonation:-** AI tools can mimic writing styles, voices, and appearances, making it easier for malicious actors to impersonate others. A fraudster could, for instance, use an AI-generated voice to scam a victim's contacts or publish false information that appears to come directly from the victim.
  5. **Permanent Digital Footprints:-** AI enhances the ability to collect, analyze, and store online activity. This makes it easier for employers, governments, or adversaries to uncover and weaponize past statements, even if those statements were made in different contexts or under different circumstances. AI's ability to aggregate and analyze historical data can unfairly tarnish reputations.
  6. **Invasive Surveillance and Doxxing:-** AI-powered surveillance tools can track online and offline behavior, often without consent. These tools, combined with AI's ability to predict patterns or analyze connections, increase the risks of doxxing—revealing personal information publicly. Such incidents can lead to harassment, reputational damage, and even physical harm.

After thorough examination and study of numerous cases on these topics, the court resorted to the principle of reconciling conflicting rights and embraced the concept of balancing these rights. Regarding the legality of criminal defamation, the Court noted that it could be evaluated based on the principles of unity and duty as outlined in the Constitution's Preamble. Yet, the Court struggled to determine definitively whether criminal defamation was inherently harmful to freedom of speech and expression, but ultimately decided that it did not contravene any constitutional articles and therefore its presence did not represent an unreasonable limitation<sup>12</sup>.

---

<sup>12</sup> **Leslie Y. Garfield Tenzer**, *Defamation in the Age of Artificial Intelligence* (Aug. 18, 2023), available at <https://ssrn.com/abstract=4545070> or <http://dx.doi.org/10.2139/ssrn.4545070>.

## 8. THE SOCIETAL IMPACT OF CYBER DEFAMATION IN THE AI ERA

The harm caused by AI-facilitated cyber defamation is far-reaching, affecting individuals, institutions, and societies at large:

- **On Individuals:** Reputational damage in the digital age can lead to profound personal consequences, including emotional distress, career setbacks, and social ostracism. Once defamatory content is online, it often persists indefinitely, perpetuating harm over time.
- **On Businesses and Institutions:** Companies and public institutions are increasingly vulnerable to AI-driven smear campaigns that undermine trust, damage brand reputations, and disrupt operations.
- **On Democratic Discourse:** Cyber defamation undermines public trust in information systems, erodes constructive dialogue, and weakens democratic accountability by spreading misinformation and targeting dissenters<sup>13</sup>.

## 9. THE CRITICAL IMPORTANCE OF A ROBUST LEGAL FRAMEWORK FOR AI-DRIVEN DEFAMATION CONTENT

As AI becomes central to moderating and mitigating defamation in the digital age, its deployment raises significant legal and ethical questions. A robust legal framework is crucial to ensure that AI technologies are used responsibly, effectively, and fairly while safeguarding individual rights such as freedom of speech and privacy. Without clear legal structures, AI-driven defamation moderation could lead to inconsistent enforcement, censorship, or misuse, exacerbating societal harm rather than mitigating it for the following reasons.

1. **Accountability and Governance** AI systems operate largely autonomously, making it difficult to pinpoint responsibility for errors, such as false accusations of defamation or failure to detect harmful content. A legal framework ensures that accountability is distributed among AI developers, platform operators, and content moderators. Laws like the EU's Digital Services Act (DSA) impose specific obligations on platforms to use AI responsibly, ensuring transparency and liability for harmful outcomes.
2. **Protecting Freedom of Speech:** AI systems may overreach, flagging legitimate opinions or critiques as defamatory content. Legal standards must define clear boundaries for defamation to avoid censorship and uphold freedom of expression. Balance protecting reputations with the right to public discourse and dissent. The U.S.

legal concept of “actual malice” in defamation cases can guide AI systems to differentiate intentional harm from valid criticism.

- 3. Cross-Border Challenges:-** Defamation laws vary widely across jurisdictions. Content flagged as defamatory in one country might be permissible in another. A global legal framework or cooperative agreements to address cross-border defamation cases involving AI moderation. The Budapest Convention on Cybercrime provides a model for international cooperation in online crimes, including defamation.
- 4. Transparency and Ethical Oversight:-** Many AI algorithms used to moderate content are "black boxes," making their decision-making opaque. Legal frameworks must mandate transparency in how AI identifies, flags, and removes content. Require AI systems to provide explainable decisions and allow users to appeal automated takedowns.
- 5. Privacy Protection:-** AI systems often analyze personal data to detect defamatory content, potentially infringing on user privacy. A legal framework must ensure compliance with privacy laws such as GDPR and define how personal data can be processed by AI. Protect user data while enabling effective moderation.
- 6. Standardizing AI Implementation:-** Without legal guidelines, AI implementation varies widely across platforms, leading to inconsistent enforcement of defamation policies. A standardized legal framework to ensure uniformity in how AI is developed and deployed for defamation detection, which shall establish minimum technical and ethical standards for AI tools.

#### **Case Study: The EU Digital Services Act (DSA)**

The DSA, implemented by the European Union, serves as an example of a comprehensive legal framework addressing online content moderation in AI Regulation and Transparency Obligations. This framework showcases how legal systems can regulate AI-driven defamation while maintaining user rights and platform accountability. A robust legal framework is indispensable for ensuring the responsible use of AI in moderating defamatory content. It provides clarity, accountability, and safeguards individual rights, creating a balanced approach to managing defamation in the digital age. Such frameworks must evolve alongside AI technologies to remain effective, fostering trust in both AI systems and the platforms that deploy them. By aligning legal, ethical, and technological standards, society can mitigate the harm caused by defamation while preserving the principles of free expression and privacy.

## 10. EFFECTIVE AI IMPLEMENTATION TO MITIGATE SOCIETAL HARM CAUSED BY DEFAMATION IN THE DIGITAL AGE

The digital age has dramatically increased the speed and scale at which information, including defamatory content, spreads online. Traditional methods to address defamation—such as manual content monitoring or legal interventions—struggle to keep pace with this rapid dissemination. Artificial intelligence (AI) offers scalable, efficient, and innovative solutions to identify, manage, and mitigate the societal harm caused by defamation in this environment. However, its implementation must be carefully designed to balance effectiveness, ethics, and accountability in the following ways.

- 1. Real-Time Content Monitoring:** AI systems can process vast volumes of data across social media, news sites, and blogs to identify defamatory content. Advanced natural language processing (NLP) algorithms can detect false or malicious statements targeting individuals or organizations and recognize patterns associated with defamatory speech, such as derogatory language or inflammatory tone. For instance, platforms like Facebook and Twitter use AI-driven tools to flag potentially harmful posts for human moderators to review.
- 2. Fact-Checking and Misinformation Detection** Defamation often arises from the spread of false information. AI systems can cross-check statements against reliable databases or fact-checking repositories to identify inaccuracies or fabrications and flag content that appears misleading or defamatory. Tools like Google's AI-based fact-check algorithms or platforms like ChatGPT and Bard are being explored to validate claims.
- 3. Sentiment and Context Analysis** AI tools equipped with sentiment analysis can assess the emotional tone of content to distinguish between malicious defamation and constructive criticism. Contextual AI can further analyze the intent behind a statement and broader conversation in which the content is embedded, with this capability reducing false positives, ensuring that legitimate free speech is not wrongly penalized.
- 4. Automated Content Moderation** AI can automate the removal of defamatory content at scale. When combined with machine learning models trained on examples of defamatory speech, AI can identify and take down harmful posts before they cause widespread damage. Notify platform users about why their content was removed, promoting awareness and deterrence. Companies like YouTube and TikTok use AI moderation tools to handle millions of user submissions daily.

5. **Proactive Monitoring Through Predictive Analytics** Predictive AI models can analyze patterns of online behavior to anticipate potential defamatory activity. By identifying high-risk users or groups prone to harmful behavior, AI can help prevent the spread of defamatory content and enable platform administrators to intervene early.
6. **Scalability and Multilingual Support** Unlike human reviewers, AI can operate 24/7 and handle content in multiple languages simultaneously. This is crucial for global platforms where defamatory content may originate in various linguistic and cultural contexts. AI models trained on diverse datasets can address such challenges effectively.

## 11. BEST PRACTICES FOR EFFECTIVE AI IMPLEMENTATION

- i. **Human-AI Collaboration:-** AI systems should complement, not replace, human moderators. Human review ensures that nuanced cases are handled thoughtfully.
- ii. **Transparency and User Education:-** Platforms should clearly explain how AI systems identify and manage defamatory content, and educating users about these systems can foster trust and reduce resistance to content moderation.
- iii. **Regular Updates and Feedback Loops** AI models must be regularly updated to adapt to new forms of defamation, slang, and cultural shifts, and incorporating user feedback can improve system accuracy and fairness.
- iv. **Partnerships and Shared Databases** Collaboration between platforms, fact-checking organizations, and AI developers can enhance the efficiency of defamation detection, and Shared databases of flagged content can improve AI training and reduce redundancy.
- v. **Ethical Guidelines and Governance:-** Establishing clear guidelines for AI development and deployment ensures that systems operate within legal and ethical boundaries.

Effective AI implementation has the potential to revolutionize the fight against defamation in the digital age. By enabling real-time monitoring, fact-checking, and predictive analytics, AI can significantly reduce the societal harm caused by defamatory content. However, to ensure its success, platforms must address challenges such as bias, transparency, and ethical concerns. A balanced approach that integrates robust AI technologies with human oversight and legal safeguards can create a safer and more responsible online environment.

The challenges of AI and cyber defamation demand a multi-stakeholder response involving governments, the judiciary, technology companies, civil society, and individuals. Each has a

distinct role to play:

- **Governments and Regulators:** Establish clear, enforceable standards that balance innovation and accountability.
- **Judiciary:** Interpret laws with an understanding of new technologies, ensuring fair and proportional remedies for reputational harm while protecting free expression.
- **Tech Platforms:** Invest in AI tools for moderation, ensure transparency in decision-making, and collaborate with regulators and civil society to address misuse.
- **Civil Society:** Advocate for ethical practices, hold platforms accountable, and provide support for victims of cyber defamation.
- **Individuals:** Engage responsibly online, report defamatory content, and support efforts to foster respectful and constructive discourse.

## 12. CONCLUSION

The digital era has transformed the ways we communicate, empowering individuals to share their thoughts with global audiences in an instant. However, this freedom has also introduced challenges, especially in balancing two fundamental rights: the right to free speech and the right to reputation. Both are cornerstones of a democratic society, but the rise of cyber defamation has shown how the abuse of one can severely harm the other.

Artificial intelligence has emerged as a powerful tool to address the scale and complexity of defamation in the digital age. From identifying harmful content to enabling real-time moderation, AI offers immense potential. However, it also brings risks—such as over-censorship, algorithmic bias, and the suppression of legitimate discourse—underscoring the need for careful governance.

AI represents a double-edged sword for free expression and online reputations. On one hand, it democratizes creativity, amplifies marginalized voices, and fosters global communication. On the other hand, it introduces new risks, including deepfakes, misinformation, and invasive surveillance, that can irreparably harm reputations. By fostering ethical development, robust regulations, and public education, society can leverage AI's potential to enhance free expression while protecting individuals from reputational harm. Balancing these forces is critical to ensuring that AI remains a tool for empowerment rather than destruction.

A robust legal framework is critical to navigating this complex landscape. Such a framework

must not only guide the responsible use of AI but also ensure that individuals are protected against defamatory harm without stifling their ability to express opinions. Transparency, accountability, and a commitment to ethical standards must underpin the integration of AI into content moderation.

Ultimately, the balance between free speech and the right to reputation is not a zero-sum game. Thoughtful regulation, coupled with technological innovation and human oversight, can create a digital environment where both rights coexist harmoniously. By fostering collaboration among legal systems, technology providers, and civil society, we can uphold the principles of justice, fairness, and dignity in the ever-evolving digital era.

The AI-driven age brings immense potential for innovation and expression, but it also presents significant challenges to the preservation of individual dignity and societal trust. Combating the threat of cyber defamation in this era requires a proactive and collaborative approach that combines legal safeguards, ethical AI development, and robust public awareness.

The goal must be to foster a digital ecosystem where free expression flourishes alongside respect for reputational rights. By advancing transparent laws, responsible technological practices, and informed societal behavior, we can mitigate the risks of AI-driven harm while harnessing its benefits to create a safer, more equitable online environment. Ultimately, the balance we strike today will define the integrity and inclusivity of our digital future<sup>14</sup>.

---

<sup>14</sup> Swarnali Das, *Right to Reputation in the Age of Artificial Intelligence: Navigating the Challenges of Synthetic Media*, 6 INT'L J. LEGAL SCI. & INNOVATION 183 (2024), <https://ijlsi.com/wp-content/uploads/Right-to-Reputation-in-the-Age-of-Artificial-Intelligence.pdf>.