DLC Image: Pratibimbo
ISSN: 2583-8881
https://image.dispurlawcollege.org

Generative Artificial Intelligence and Interpersonal Cybercrime: A Human Rights Perspective

Himashree Konwar
Research Scholar, National Law University and Judicial Academy, Assam
Email: himashreekonwar94@gmail.com

Abstract

Generative Artificial Intelligence is much more efficient at producing realistic content from vast existing datasets and has already sparked a technological evolution. With such advancements, the world has also seen a concerning rise in cybercrime activities. The internet has witnessed several new forms of interpersonal cybercrime, where the misuse of Generative Artificial Intelligence has enabled individuals to be targeted, defamed, and harmed mentally and physically through the threat of uploading manipulated content online. This emerging menace of cybercrime not only jeopardises individuals' security but also poses serious threats to fundamental human rights. In this paper, the author aims to analyse the human rights aspects of crimes affecting individuals in cyberspace and the legal landscape surrounding these growing concerns. The author limits her study to individual harm only; no financial cybercrime offences are incorporated in the study. This paper critically examines international human rights instruments, relevant constitutional protections, and statutory frameworks governing cybercrime that impact individual rights. The research finds that existing legal frameworks are technologically neutral, often insufficient to address the speed, anonymity, and scale of such abuse. The paper advocates for a rights-based regulatory approach that incorporates victim-centric remedies and AI-specific safeguards to ensure protection in the digital age.

Keywords- *Technology, Generative Artificial Intelligence, Interpersonal Cybercrime, Human Rights, Cyberspace*

1. Introduction

Generative Artificial Intelligence is considered one of the most efficient technologies humans have ever innovated. The significant evolution of Generative Artificial Intelligence in digital technologies has sparked a worldwide evolution.[1] It is an Artificial Intelligence tool capable of producing hyper-realistic text, images, audio, and videos. Unlike other tools of AI, generative AI possesses the capacity to create synthetic content that closely resembles human generated-outputs. The super-efficient features of generative AI have greatly impacted human

[1] Asifa Hassan, "Cybercrimes: Emerging trends, technologies and legal aspects in India" 5(1) *International Journal of Criminal, Common and Statutory Law* 212 (2025).

communication and interaction. It has transformed digital communication, content creation, and online interaction globally.[2] However, such efficient technologies are a blessing until they fall into the wrong hands. The exploitation of generative AI features may intensify criminal activities in cyberspace.

The scope of cybercrime is no longer limited to machines nowadays. Likewise, communication between people has also reached far beyond physical meetings and spoken words. Increasing digital dependency and evolving digital engagement have built a fertile ground for interpersonal cybercrime. Incorporation of generative AI in digital communication and content creation has enabled new forms of interpersonal cybercrime, blurring the boundaries between innovation and exploitation. With the advent of generative AI and unprecedented global connectivity, sophisticated cyberattacks against individuals have emerged. Cyber miscreants have become well-equipped to commit crimes in cyberspace that target individuals by manipulating such super-efficient technologies.[3]

In India, since the adoption of generative AI in communication, interaction, and content creation, a significantly accelerated rate of interpersonal cybercrime has been noticed. The intersection of generative AI and interpersonal cybercrime raises urgent concerns. In the rapidly growing digital ecosystem, the vast user base across social media and other digital platforms is particularly vulnerable to these developments.

2. Understanding Generative Artificial Intelligence

Generative AI is a branch of Artificial Intelligence that creates new original content, including text, images, code, music, and videos, by learning patterns from existing data.[4] Generative AI relies on advanced technologies like machine learning models called deep learning, and models such as Large Language Models, Generative Adversarial Networks. Such technologies stimulate the learning and decision-making process of the human brain. These models work by identifying and encoding the patterns and relationships in huge amounts of data, and then using that information to understand users' natural language requests or questions and respond with relevant new content. By incorporating advanced technologies, generative AI enables digital systems to generate human-like responses and realistic content. It has become far more efficient in creating hyper-realistic content, which is highly convincing as real. Therefore, generative AI has brought a wave in the digital ecosystem in recent times.

2.1 Generative AI and Its Tools

Generative AI tools are software applications built on these technologies that allow users to generate content through simple prompts or instructions. Text generation tools generate human-like text for writing, research, coding or conversation. ChatGPT, Google Gemini, Claude AI and Jasper are some of the text generation tools. Likewise, DALL-E, Midjourney, Stable Diffusion, and Adobe Firefly are image generation tools which generate super-realistic images from a textual prompt. Similarly, Runway ML, Synthesia, and Pictory AI are video generation tools which create videos using AI-generated visuals and scripts. Moreover, ElevenLabs, Murf

[2] R. Shital, K. Swapna, "Cyber Crimes in India: A Critical Review" 12(4) *Journal of Emerging Technologies and Innovative Research* 582 (2025).

[3] Elena Shalevska, "Human Rights in the Age of AI: Understanding the Risks, Ethical Dilemmas, and the Role of Education in Mitigating Threats" 1(2) *Journal of Legal Political Education* 40 (2024).

[4] IEEE Computer Society Digital Library, *Available at*: https://www.computer.org/csdl/proceedings-article/eisic/2013/06657122/12OmNBTs7A9. (Last visited on February 25, 2026).

AI, and AIVA are voice and audio generation tools that generate realistic synthetic voices or music.

These generative AI tools have numerous beneficial applications in several fields of human life. Such tools have already been adopted and are being used in individuals' daily lives. However, they also pose risks when misused, particularly in the context of interpersonal cybercrime. In creating morphed and manipulated content, such tools of generative AI are being sensitively used. In making contents of deepfakes, pornography, impersonation, online harassment, and identity fraud, these advanced tools are misused by the miscreants.[5]

3. Understanding Interpersonal Cybercrime

With the introduction of advanced technologies, several new challenges have arisen in cyberspace in recent decades. Interpersonal cybercrime is one of such threats that has been surging over time. In interpersonal cybercrime, the miscreants commit harm against individuals with whom they interact, communicate, or share some sort of real or imaginary affection. It is a form of cybercrime activity that specifically targets individuals.[6] When cybercrime perpetrators attack people on digital sites, whether known or unknown, to accomplish their malicious intentions, it is considered an interpersonal cybercrime. In many cases, victims are already known to the perpetrators. They target individuals to harm their social reputation, mental well-being, and to commit blackmail, emotional distress, sextortion, etc.

The activities of interpersonal cybercrime can be further categorised into three types: online harassment and bullying, sexual exploitation and abuse, and financial and identity crime. However, the author in this paper focuses on the first two categories, namely online harassment and bullying, and sexual exploitation and abuse. These issues have become a major concern as the incidence of these two categories has surged in the digital space recently.

3.1 Shifting Paradigms of Interpersonal Harms in the Generative AI Era

The incorporation of Generative AI in communication and interaction has brought a drastic change in the scope of criminal activities in the digital space. By using the super-efficient features of Generative AI, cyber miscreants can now commit more intensified cyber-attacks targeting individuals.[7] In recent times, the world has witnessed several new forms of cyber incidents where miscreants have misused Generative AI in creating manipulative content. They can morph any photo or video of an individual with just one click and can use it against that person. The features of voice alteration can be misused in altering voice of any individual and use it in committing interpersonal harm. Moreover, the tools of Generative AI for creating text content, i.e. Large Language Models, can also be used in manipulative text, emails, and code-based assistance to target individuals in cyberspace. The miscreants, with the help of such features of Generative AI, effortlessly make so real content that it is impossible to differentiate. As a result, the cyber miscreants have become more efficient in committing sophisticated

[5] Divyani Varma, Dr. Richa Shrivastava, Dr. Niti Nipuna, Dr. Vishal Sharma, "The Rising Threat of Cybercrimes in India: Challenges, Legal Provisions, and Solutions" 4(1) *International Journal of Contemporary Research in Multidisciplinary* 180 (2025).

[6] Key Issues, Module 12: Interpersonal Cybercrime, E4J University Module Series, UNODC, Doha Declaration: Promoting a Culture of Lawfulness *available at*: https://www.unodc.org/e4j/en/cybercrime/module-1/keyissues/cybercrime-in-brief.html (last visited on February 26, 2026).

[7] Himanshu Jagwani, Khushbu Jagwani, and Sukhjeet Kaur Matharu, "The use of Generative AI by Financial Cybercriminals: Emerging Threats and Countermeasures for Banks" 8(2) *International Journal of Financial Management and Economics* 1307 (2025).

cyber-attacks against individuals. Such activities are highly sensitive and challenge individuals' autonomy in digital space.

Over the years, the world has witnessed an increasing rate of content in which Generative AI is being misused to cause interpersonal harm in cyberspace. With such AI-generated content, individuals are easily targeted, manipulated and misled. Many new patterns of cybercrime against individuals have evolved in recent years, posing an alarming threat to the cyber ecosystem.

A few new dimensions have emerged over time with the adoption of generative AI in communication and interaction between individuals. We can summarise such offences under categories

3.2 Online Harassment and Bullying

It refers to acts where a wrongdoer intentionally, using electronic communication, harms, threatens, humiliates another person and causes physical or mental distress. Such offences are still developing in cyberspace; however, criminal activities such as cyber stalking, cyberbullying, defamation, and hate speech are some instances where victims endure such suffering.

3.2.1 Cyber Stalking- Cyber stalking is a pattern of activities where a person with malicious intentions, by using electronic communication, harasses another person. The perpetrators in cyber stalking consistently follow the targeted individual on digital platforms, threaten and harass them by sending unwanted emails, texts, images, voice messages, etc.

3.2.2 Cyber Bullying- Cyberbullying is a form of crime where individuals can be harassed using digital space by humiliating others through spreading negativity repeatedly, spreading harmful and embarrassing content.[8] Impersonation, doxing, flaming, trolling, and sharing non-consensual images are some of the forms of acts that can be referred to as bullying in cyberspace.

3.2.3 Cyber Defamation- Cyber defamation is the act of spreading false and damaging statements on digital platforms with the intention of harming someone’s reputation. It can be in the form of false spoken statements in video or audio, or false written statements online.

Online Hate speech- Online hate speech is considered spreading any form of expression on digital platforms that attacks, demeans or uses pejorative or discriminatory language with reference to a person or a group based on their race, religion, ethnicity, gender, sexual orientation, disability, political opinion, etc.

3.3 Sexual Exploitation and Abuse

The activities of sexual exploitation and abuse are also shifting to the digital space. Nowadays, an individual can be exploited and abused even without any touch. Such activities refer to a range of offences that use information and communication technologies to facilitate the sexual exploitation and abuse of victims. Image-based sexual abuses such as cyber voyeurism, sextortion, child pornography, online child abuse, deep fake, etc., are some of the instances of acts where an individual, irrespective of any age or gender, can be exploited and abused on the internet.

[8] Autumn Slaughter, Elana Newman, “New Frontiers: Moving Beyond Cyberbullying to Define Online Harassment” 1 *Journal of Online Trust and Safety* 26 (2022).

3.3.1 Cyber Voyeurism- Cyber voyeurism is an act of secretly observing, recording, or disseminating private acts with malicious intention by using digital technologies. Such activities involve the unauthorised capture or sharing of images of individuals in private situations where privacy is reasonably expected.

3.3.2 Online sextortion- The term sextortion blend denotes the blend of 'sex' and 'extortion', which means when a perpetrator uses explicit or intimate sexual images or videos of a victim to coerce them into providing more sexual content or sexual favours. In such activities, the perpetrators first gain the trust of the individual, and later they start threatening or blackmailing.

3.3.3 Online child abuse and child Pornography- Activities of online child abuse and child pornography are a concerning offence, where children who are sexually assaulted can be affected drastically. It refers to the activities of creating any visual depiction of sexually explicit conduct involving a person below the age of 18 years.[9] Engaging a person in such sexual activities and spreading it over an online platform raises a serious threat to innocent children.

3.3.4 Deep Fake- Deep fake refers to the morphing of any picture or video, or any content of any person, with malicious intention to defame or to sexually expose that person. Deep fake content is so perfectly made that people can be easily manipulated and convinced that the content is real.

3.4 Vulnerability to Interpersonal Cybercrime Harms

In the age of Generative AI, every segment of society is vulnerable to interpersonal cybercrime and violence. However, some sections of society are considered more vulnerable to such criminal activities, i.e. women, children, adolescents, elderly persons, the third gender community and physically disabled persons. These people are easily targeted in interpersonal cybercrime offences.

4. Human Rights Implications against Interpersonal Cybercrime

4.1 Right to privacy

In the digital ecosystem, privacy rights are highly vulnerable. The development of technologies such as generative AI and its tools, which can generate any content by analysing input data, significantly heightens concerns about individuals' privacy in cyberspace.[10] Platforms such as social media and other communication channels have greatly increased opportunities for people to share personal and sensitive information.[11] Cybercriminals, by misusing generative AI technologies, can easily create deepfakes, manipulated, and altered content that harms individuals' privacy rights. Unauthorised generative AI-generated representations infringe personal autonomy and control over one's identity.

4.2 Right to dignity and reputation

Cybercriminals nowadays create content using generative AI. Such content is super realistic, so that any person can be impersonated on any digital platform. Miscreants are causing harm

[9] Laurie S. Ramiro, Andrea B. Martinez, Janelle Rose D. Tan, Kachela Mariano, Gaea Marelle J. Miranda, Greggy Bautista, "Online Child Sexual Exploitation and Abuse: A Community Diagnosis Using the Social Norms Theory" 96 *Child Abuse and Neglect* 104080 (2019).

[10] Dr. Aradhana Parmar, "Laws to Protect Right to Privacy in the Internet Era" 11 *Journal of Advances and Scholarly Research in Allied Education* 341 (2016).

[11] Swati Sharma, Vikash Kumar Sharma "Cybercrime Analysis on Social Media" 11(1) *Journal of Computer* 7 (2020).

by making and spreading such content on social media and other platforms.[12] Such activities are highly injurious to the dignity and reputation of individuals. Sometimes, people are being blackmailed by the threat of spreading such content. With the advancement of such technologies, which create realistic content, the fear of damage to reputation and dignity has intensified recently.

4.3 Freedom of Expression

The rights of individuals' freedom are also at great stake in the digital world. Offences of spreading hate speech on social media and other criminal activities of spreading defamatory AI- generated content can often be seen where generative AI is used by the cybercrime perpetrators. Such misuse of technology can significantly damage an individual's reputation and credibility, often leading victims to withdraw from online platforms to avoid further harassment. Consequently, such activities on digital platforms significantly undermine the effective realisation of freedom of expression.

5. International Instruments on Interpersonal Cybercrime Implying Generative AI

International human rights instruments have evolved recently to secure the rights against the AI-generated content that harms individual liberty to a great extent. These instruments advocate for transparency, accountability, and the protection of dignity, privacy and freedom of expression in cyberspace.

5.1 United Nations Organisation and its Declarations

UNESCO's recommendation on the Ethics of Artificial Intelligence was first initiated as a global standard-setting instrument in 2021, which has been adopted by 194 member states.[13] It advocates for mandatory ethical impact assessment for high-risk AI, prohibits AI systems for social scoring and mass surveillance, and ensures AI decisions are consistent and transparent.

In March 2024, the United Nations adopted a declaration that has been ratified by 193 member states, which emphasises the recognition of human rights in AI-generated content and calls for AI to be safe, secure, and trustworthy.[14]

In September 2024, a Global Digital Compact was adopted by 193 member states, which seeks to bridge the digital divide, ensure human-centric AI, and promote responsible innovation, establishing a multi-stakeholder model to guide future technology policies globally.

The United Nations Human Rights Office of the High Commissioner also initiated the Business and Human Rights in Technology Project in 2019, which guides the application of the United Nations Guiding Principles to the tech sector, focusing on addressing human rights risks in digital technologies and business models.[15]

[12] Divya Hariharan, Harsh Kumar Singh "AI, Data Privacy, and Right to be forgotten: Navigating Human Rights in the Age of Generative Technology" 5(2) *Indian Journal of Integrated Research in Law* 2151 (2024).

[13] UNESCO, *available at*: https://www.unesco.org/en/artificial-intelligence/recommendation-ethics (last visited on February 28, 2026).

[14] UNODC, *available at*: https://www.unodc.org/e4j/en/cybercrime/module-9/key-issues/assets--vulnerabilities-and-threats.html (last visited on February 28, 2026).

[15] United Nations, *available at*: https://news.un.org/en/story/2024/03/1147831 (last visited March 1, 2026).

To secure freedom of expression from generative AI-driven manipulation and disinformation, the United Nations in the year 2025 launched a Joint Declaration on AI, Freedom of Expression and Media.[16]

5.2 Role of the European Union

The European Union has been playing an active role in combating the AI-driven violence in cyberspace. In March 2024, the Union adopted the first legally binding international treaty, which has been designed to ensure AI systems where human rights, democracy and rule of law shall be secured.[17] Another significant development is the European Union AI Act, 2024, which has been adopted for setting generative AI rules, transparency obligations, and prohibited practices. Moreover, the EU Digital Services Act, 2024, has been enacted, and it enforces stricter rules on content moderation, transparency and safety. It ensures heavy fines for non-compliance for intermediaries like social media, marketplaces and application stores. Additionally, the EU Directive on Combating Violence against Women, adopted in 2024, is also considered an important step in securing human rights for women in cyberspace. It criminalises the non-consensual sharing of intimate, AI- driven material such as deepfakes and ensures that such activities are considered as aggravating circumstances.

6. Indian Legal Framework in Ensuring Human Rights in the Age of Generative AI

6.1 Constitutional Measures

The Constitution of India provides the basic human rights necessary to live a dignified life as a human being in a digital ecosystem.[18] The "golden triangle" comprising Articles 14, 19 and 21 provides a foundational framework to protect citizens from AI- driven manipulated content that harms individuals' rights to a great extent. Article 14 ensures equality before the law and equal protection of the law.[19] This provision protects the rights of citizens against any discriminatory and arbitrary acts using AI- driven content moderation or decision-making features. Article 15 ensures rights against any discrimination based on religion, race, caste, sex, or place of birth by AI- manipulated content, which seriously causes reputational, emotional and financial damage to individuals. Moreover, the right to freedom of speech and expression under Article 19 (1) (a) incorporates online content, including AI- generated material. AI-manipulated content also carries the scope of interference in the privacy rights of individuals. The right to privacy is protected under Article 21 and recognised in a landmark judgement, *Justice K.S. Puttaswamy v. Union of India*[20], which also extends to AI-driven content involving massive data collection, profiling, or surveillance, and must adhere to the principles of legality, necessity, and proportionality.

---

[16] UNODC, *available at*: https://www.unodc.org/e4j/en/cybercrime/module-1/key-issues/cybercrime-in-brief.html (last visited on February 27, 2026).

[17] European Union, *available at*: https://commission.europa.eu/news-and-media/news/digital-services-act-keeping-us-safe-online-2025-09-22_en, (last visited on March 2, 2026).

[18] M.P. Jain, *Indian Constitutional Law* 102 (LexisNexis, India, 8th Edn., 2018).

[19] Prof. (Dr.) Ashok Kumar Rai, Dr. Santosh Kumar, "Constitutional Rights in the Digital Age: The Intersection of Constitutional Rights and Digital Technologies" 1(3) *International Journal of Multidisciplinary Research* 55 (2023).

[20] AIR 2017 SC 4161.

6.2 Legislative Frameworks

In recent years, the Indian legislature has taken several initiatives to tackle the AI-generated individual harm. The legislature has focused on more structured, binding regulations to deal with AI-generated content, misinformation, and algorithmic bias. The Information Technology Act, 2000, was enacted to specifically deal with cyber offences. From time to time, it has been amended to incorporate the new criminal activities arising with the emergence of technologies. In 2025, India's Ministry of Electronics and Information Technology unveiled the Seven Sutra Framework as part of its AI governance guidelines, based on a principle-based "light touch" approach designed to promote responsible and safe AI adoption.[21] The seven core principles are trust is the foundation, people first, innovation over restraint, fairness and equity, accountability, understandable by design, and safety, resilience & sustainability. Very recently, the Information Technology (Intermediary Guidelines) Amendment Rules, 2026, have been introduced to target synthetically generated information and mandate social media and other platforms to label AI-generated content. This Rule also emphasises a 3-hour takedown window for reported deepfakes or illegal content that harms individuals' reputation, emotional and financial well-being. Moreover, the Digital Personal Data Protection Act, 2023, also incorporates provisions for handling personal data in the AI-generated system, focusing on consent, accountability and purpose-limited. Additionally, AI-driven cyber offences also come under the provisions of Bharatiya Nyaya Sanhita, 2023, which has replaced the Indian Penal Code. This new milestone in criminal law serves as a legal tool for prosecuting the misuse of AI.

6.3 Judicial Stance on Securing Individuals' Rights in Cyberspace

The judiciary has taken the matter of individuals' harm in cyberspace on a serious note. The judiciary in deciding the matters related to AI- generated cybercrime harms has adopted a proactive-defensive approach, utilising existing laws to combat novel harms. The courts have, from time to time, held intermediaries accountable for failure to swiftly remove harmful non-consensual AI-generated content such as deepfakes, which cause serious effects on individuals. The courts have rigorously emphasised securing basic human rights in cyberspace. The Supreme Court in *Subramaniam Swamy v. Union of India*[22], upheld the constitutional validity of criminal defamation laws incorporated under Sections 499 and 500 of the Indian Penal Code. The Apex court ruled that the right to reputation and dignity is a fundamental right, and thus, criminal defamation is a reasonable restriction on freedom of speech.

Again, in *Justice K.S. Puttaswamy v. Union of India*[23], the Supreme Court already in the year 2017 recognised the right to privacy as a fundamental right under Article 21 of the Constitution of India. The Supreme Court ruled that privacy is an intrinsic part of individual dignity, liberty and autonomy, which set a benchmark for data protection and state surveillance.

As digital space has become an inseparable part of human life recently, such fundamental human rights shall be protected even in cyberspace. Without preserving the dignity, privacy and personal rights of individuals, the digital space would become a chaos.

[21] PIB Headquarters, *available at*:

https://www.pib.gov.in/PressReleasePage.aspx?PRID=2228315&reg=3&lang=2, (last visited on March 3, 2026).

[22] AIR 2016 SC 2728.

[23] *Supra* note 10 at 17.

7. A Way towards Right-Based Regulatory Framework

In the digital era, the concept of fundamental human rights in cyberspace has become very complex. As society progresses with digitalisation, some new phenomena of interpersonal harm are evolving which threaten individuals' privacy, reputation, dignity and well-being.[24] The cyber legal framework should focus on securing human rights as well in this digital era, while regulating digital activities rather than just prioritising security or state control. Such an approach emphasises that users are the subjects of rights, not just objects of regulations. In the age of AI-generated content, a human rights-oriented framework should incorporate mandatory transparency obligations, criminalisation of malicious activities, strengthening data protection standards, accessible grievance redressal mechanisms for victims, platform accountability, and content moderation duties.[25]

8. Conclusion

The transition of digital technologies to generative AI signifies a transformative shift in the nature, scale and intensity of interpersonal cybercrime. AI-enabled harms such as deepfake-based sexual exploitation, automated harassment, identity fabrication and synthetic defamation threaten core human rights, including privacy, dignity, reputation, equality and psychological integrity. While the international human rights norms and constitutional mandates guarantee the basic fundamental rights to individuals and the legislative statutes provide a normative foundation for protection, no specifically designed legal instrument to tackle the unique characteristics of generative AI, particularly its super-efficient features to generate hyper-realistic content, anonymity, scalability and cross-border dissemination. Therefore, a rights-based regulatory framework is essential to govern the AI-driven cyber offences that specifically threaten human rights in cyberspace. An approach that integrates principles of proportionality, transparency, due diligence, and platform responsibility while preserving legitimate freedom of expression. The prospective generative AI governance must be anchored in human dignity and constitutional morality. Legal reform, judicial interpretative innovation, and ethical AI design must converge to ensure that technological advancement does not erode fundamental human rights. Safeguarding individuals in the digital ecosystem is not merely a regulatory necessity but a democratic imperative.

[24] Dave Lewis, Joss Moorkens, "A Rights-Based Approach to Trustworthy AI in Social Media" 1(13) *Social Media and Society* 09 (2019).

[25] Tariq Rahim Soomro, Mumtaz Hussain, "Social Media- Related Cybercrimes and Techniques for their Prevention 24"(1) *Siendo* 307 2019.