

SHODH SAMAGAM

ISSN : 2581-6918 (Online), 2582-1792 (PRINT)



Artificial Intelligence, Social Media and Cyber Stalking: Emerging Threats to Human Rights

Udit Agnihotri, Ph.D., Faculty of Law
Major SD Singh University, Farrukhabad, Uttar Pradesh, INDIA

ORIGINAL ARTICLE



Author

Udit Agnihotri, Ph.D.

E-mail : uditlm@gmail.com

shodhsamagam1@gmail.com

Received on : 09/05/2026
Revised on : 10/07/2026
Accepted on : 19/07/2026
Overall Similarity : 08% on 11/07/2026



Plagiarism Checker X - Report

Originality Assessment

8%

Overall Similarity

Date: Jul 11, 2026 (02:30 PM)
Matches: 186 / 2268 words
Sources: 14

Remarks: Low similarity detected, consider making necessary changes if needed.

Verify Report:
Scan this QR Code



ABSTRACT

The fusion of artificial intelligence with social media platforms has transformed cyber stalking from a nuisance into a mechanised assault on human dignity, privacy, and participatory rights. Using Indian legal and international human rights instruments as the backbone, this article maps the AI enabled techniques predictive surveillance, deepfakes, bot driven mobbing that stalkers now deploy. It examines the injury these inflict on the right to privacy, psychological integrity, and free expression, paying special attention to gendered impacts. The piece critically evaluates Indian statutory responses such as the Information Technology Act, 2000, the Bharatiya Nyaya Sanhita, 2023, and the Digital Personal Data Protection Act, 2023, alongside international obligations under the ICCPR and CEDAW. Finally, it proposes a human rights centred regulatory framework that binds platforms to safety by design principles and ensures survivor focused remedies. Without urgent legal reform and ethical recalibration of AI systems, digital spaces will remain hostile territory for the most vulnerable.¹

KEY WORDS

Artificial Intelligence, Social Media, Cyber Stalking, Human Rights, Right to Privacy, Deepfakes.

INTRODUCTION

In 2023, a young journalist in Mumbai discovered that her face had been grafted onto pornographic videos circulating on encrypted messaging groups.² The perpetrator an anonymous account she had blocked months earlier had used open source AI tools to scrape her Instagram photos and create synthetic media that were indistinguishable from reality. Each morning, she woke to hundreds of graphic threats, many generated by chatbots that never slept.³ The ordeal, which lasted

over a year, is not exceptional. It is emblematic of a new generation of cyber stalking where artificial intelligence serves as both weapon and shield, and social media algorithms accelerate the destruction.⁴

Cyber stalking, traditionally defined as a course of conduct directed at a specific person that causes a reasonable person to fear for their safety, has always had a digital shadow.⁵ Yet the character of that shadow has changed. Where once the stalker was limited by human stamina, today's perpetrator can deploy machine learning to monitor a target's every digital move, generate thousands of personalised harassing messages per hour, and mobilise synthetic identities to destroy professional and personal reputations.⁶ Social media platforms, engineered to maximise engagement, distribute the abusive content with the same neutrality they apply to a friend's holiday photograph.⁷ This article argues that the AI social media nexus does not merely amplify the old crime; it creates a qualitatively new category of harm that current legal instruments both Indian and international are ill equipped to address. It proceeds by examining the technological anatomy of AI enabled stalking, the human rights it violates, the gender specific way it is experienced, and the regulatory architecture required to hold perpetrators and platforms accountable.⁸

The Convergence of AI and Social Media in Modern Stalking

The stalker's toolkit has undergone a silent revolution. Ten years ago, the most sophisticated digital stalker might have used a keylogger or a hidden tracking app. Today, the same objective is achieved with a few hundred rupees' worth of commercial spyware that deploys machine learning models to predict a victim's movements, emotional state, and likely response to provocation.⁹ These applications, often disguised as parental monitoring or employee wellness software, harvest real time geolocation, call logs, camera feeds, and social media activity, feeding them into algorithms that construct an intimate behavioural dossier.¹⁰ The stalker no longer needs to guess when the victim is alone or distressed; the AI notifies them.

Social media platforms operate as the primary feeding ground for this intelligence. Instagram, Facebook, X (formerly Twitter), and YouTube are built on surveillance advertising architectures that incentivise the collection and retention of vast amounts of personal data.¹¹ The "data broker" industry, legal under India's current weak privacy framework, allows anyone with a credit card to purchase datasets containing a person's location history, relationship status, religious and political leanings, and mental health indicators inferred from their browsing habits.¹² When a stalker purchases such data and cross references it with social media posts, the target becomes transparent. This intersection transforms stalking from an act of discrete observation into a state of permanent algorithmic exposure.¹³

Furthermore, the engagement maximising algorithms that sustain platform business models inadvertently abet the abuser.¹⁴ When a stalker posts a deepfake video or a defamatory accusation, the platform's recommender system may interpret the sudden spike in comments—many generated by bot accounts as a signal of authentic public interest, propelling the content into trending feeds.¹⁵

AI Enabled Cyber Stalking Techniques

1. **Predictive Surveillance and Algorithmic Profiling:** Stalkerware applications now incorporate predictive analytics that process a victim's digital exhaust tweets, likes, location check ins, Uber receipts, fitness tracker data to infer future movements and vulnerabilities.¹⁶ Unlike traditional surveillance, which is retrospective, predictive stalking is anticipatory. The stalker receives an alert that the target will likely be walking through a poorly lit park at 10 p.m. or that their sentiment analysis indicates heightened anxiety, making them susceptible to a reassuring message from a fabricated identity.¹⁷
2. **Impersonation and Synthetic Media:** Artificial intelligence has democratised the creation of deepfakes. With a handful of publicly available images, a perpetrator can generate a video that depicts the victim in sexually explicit, compromising, or criminal situations.¹⁸ The technology has progressed to the point where voice cloning requires as little as three seconds of audio, after which the stalker can generate a recording of the victim speaking any words the attacker chooses.¹⁹ Such synthetic media is then used to blackmail the victim (demanding money, sexual favours, or a return to an abusive relationship), to humiliate them publicly, or to issue fake threats in the victim's voice to their family and employers.²⁰

3. **Automated Harassment and Swarm Attacks:** Bots powered by natural language processing can sustain thousands of abusive interactions per hour, cycling through promises of love, violent fantasies, and graphic imagery in a rhythm designed to destabilise.²¹ These are not random strings of text; they are context aware messages that reference the victim's recent posts, personal photographs, and even the names of children or partners scraped from public profiles.²² The sheer volume overcomes the target's capacity to block or report, creating a digital siege that renders social media uninhabitable. Relatedly, AI coordinated swarm attacks commonly called "dogpiling" mobilise hundreds of fake accounts to converge on a single post, amplifying a hostile message until it trends.²³

Human Rights Framework: The Injuries Assessed

1. **The Right to Privacy:** In *Justice K.S. Puttaswamy (Retd.) v. Union of India*, the Supreme Court of India declared the right to privacy a fundamental right protected under Article 21 of the Constitution, encompassing informational privacy, decisional autonomy, and the right to be let alone.²⁴ AI enabled stalking violates every dimension of that guarantee. The continuous extraction of personal data through stalkerware, the non consensual creation of synthetic intimate images, and the weaponisation of personal information constitute a profound intrusion into the inner sphere that *Puttaswamy* sought to shield.²⁵ The judgment's emphasis on "privacy as a necessary condition for the meaningful exercise of other freedoms" is directly relevant: a person under constant digital surveillance cannot freely associate, speak, or develop intimate relationships.²⁶ The stalker's algorithmic gaze thus strikes at the constitutional core of personal liberty.
2. **Dignity, Psychological Integrity, and the Right to Life:** The Supreme Court of India has repeatedly read the right to life under Article 21 to include the right to live with human dignity and to be free from degrading treatment.²⁷ The psychological torture inflicted by AI stalking sustained terror, sleep deprivation, suicidal ideation can be so severe that it meets the threshold of cruel, inhuman, or degrading treatment.²⁸ When a deepfake pornographic video is circulated without consent, the victim is objectified in the most literal sense: her image becomes a commodity for public consumption.²⁹ This form of abuse attacks the very notion of the embodied self that dignity protects. The CEDAW Committee, in its General Recommendation 35, has urged states to recognise online gender based violence as a human rights violation and to provide civil and criminal remedies.³⁰
3. **Freedom of Expression and Political Participation:** The silencing effect of AI driven stalking is one of its most under appreciated consequences. Women journalists, activists, Dalit and Adivasi voices, and LGBTQ+ advocates withdraw from digital spaces after sustained attacks, self censoring or deleting their accounts.³¹ In *Shreya Singhal v. Union of India*, the Supreme Court held that the right to freedom of speech and expression on the internet is constitutionally protected under Article 19(1)(a).³² Yet when a woman politician is flooded with deepfake sexual imagery and rape threats every time she tweets about policy, her constitutional right to speak is effectively extinguished by private actors, and the state's failure to protect her speech amounts to a horizontal violation of fundamental rights.³³
4. **Intersecting Vulnerabilities and the Right to Equality:** AI stalking does not strike at random; it reinforces existing hierarchies. In India, where caste, religion, and gender intersect to produce layered vulnerabilities, a Muslim woman journalist who criticises government policy is targeted simultaneously with Islamophobic, sexist, and casteist abuse, often generated at scale by AI chatbots.³⁴ Dalit women face deepfakes that invoke caste stereotypes of sexual availability, weaponizing both gender and caste subordination.³⁵ Transgender individuals are subjected to synthetic media that out them without consent or mock their gender identity, placing them at risk of physical violence.³⁶ The right to equality under Article 14 of the Constitution and the prohibition of discrimination under Article 15 are directly implicated.

Legal and Regulatory Responses: The Indian Landscape

1. **The Information Technology Act, 2000, and Its Limitations:** The Information Technology Act, 2000 (IT Act) was India's foundational cyber legislation. Section 66E punishes the intentional capture, publication, or transmission of a private area image without consent, and Section 67 criminalises the

publication or transmission of obscene material in electronic form.³⁷ After the 2008 amendment, Section 67A specifically addressed material containing sexually explicit acts. In theory, these provisions could cover non consensual deepfake pornography. In practice, victims have faced formidable obstacles: the police lack training in identifying AI generated content, the perpetrator is often anonymous and located overseas, and the victim is frequently re-traumatised during investigation.³⁸

The IT Act's intermediary liability framework under Section 79, coupled with the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, requires platforms to observe due diligence and remove certain categories of unlawful content within strict timeframes.³⁹

2. **The Bharatiya Nyaya Sanhita, 2023:** The Bharatiya Nyaya Sanhita, 2023 (BNS), which replaced the Indian Penal Code, introduces stalking as a distinct offence under Section 78, punishable with imprisonment up to three years for a first offence and five years for subsequent offences.⁴⁰ Stalking is defined as following or contacting a person despite a clear indication of disinterest, but the provision does not explicitly mention digital or AI assisted methods.⁴¹ Sections 294 to 296 of the BNS deal with obscenity, sexual harassment, and voyeurism, yet there is no specific offence for the non consensual creation or distribution of deepfake pornography.⁴²
3. **The Digital Personal Data Protection Act, 2023:** The Digital Personal Data Protection Act, 2023 (DPDP Act) grants individuals the right to access, correct, and erase personal data, and imposes obligations on data fiduciaries to process data for lawful purposes with consent.⁴³ While the DPDP Act represents a landmark step, it does not criminalise stalkerware or the purchase of stalker oriented datasets from data brokers.⁴⁴ The definition of "harm" in the Act includes "psychological harm" and "loss of reputation," but enforcement relies on a Data Protection Board that is yet to be fully constituted and whose procedural rules remain unclear.⁴⁵
4. **A Constitutional Tort and Writ Remedy Gap:** Indian courts have developed the concept of constitutional torts against the state and, in a limited fashion, against private parties under Article 226.⁴⁶ However, obtaining an injunction against an anonymous stalker who deploys AI bots on an overseas server remains practically impossible. The Supreme Court's judgment in *Puttaswamy* emphasised that "informational privacy is a facet of the right to privacy" and that the state must create a robust data protection regime.⁴⁷

CONCLUSION

Artificial intelligence and social media have reconfigured the landscape of human interaction, but they have also opened a new frontier of human rights abuse. Cyber stalking, once a crime of passion or obsession, now operates on an industrial scale automated, anonymised, and algorithmically amplified. In India, the constitutional promise of dignity, privacy, and free expression cannot be realised when citizens are forced to live under a digital Panopticon, never knowing when their image will be stolen, their voice cloned, or their life hollowed out by an unrelenting stream of synthetic abuse. The legal architecture spanning the IT Act, the BNS, and the DPDP Act—remains fragmented, reactive, and blind to the specific horrors that AI enables. Courts have laid the jurisprudential foundation, but Parliament must build upon it with a comprehensive anti cyber stalking statute that addresses predictive surveillance, deepfake abuse, and platform responsibility as indivisible parts of a single human rights challenge.

Accountability must extend to the corporations that profit from the data and engagement that fuel stalking. Human rights due diligence, algorithmic transparency, and survivor centric design are not aspirational ideals; they are binding obligations that flow from the constitutional right to life and from international treaties that India has ratified. Technological countermeasures AI enabled detection, deepfake identification, and rapid response protocols—offer realistic hope, but only if they are embedded within a regulatory framework that prioritises the safety of the most vulnerable over the profits of the most powerful. The digital future can be one of empowerment and connection, but only if we decide, as a society, that a person's right to live free from fear extends fully into the online world. That decision must be made now; every day of legislative and corporate inaction leaves millions of Indians exposed to a regime of algorithmically driven terror that no human being should have to endure.⁴⁸

REFERENCES

1. Justice K.S. Puttaswamy (Retd.) v. Union of India (2017) 10 SCC 1, para 298; Human Rights Council, The Right to Privacy in the Digital Age: Report of the United Nations High Commissioner for Human Rights, UN Doc A/HRC/48/31 (2021) para 3.
2. This vignette is based on multiple documented accounts, see generally Amnesty International, Troll Patrol Report: Violence and Abuse against Women on Social Media (Amnesty International Ltd, 2020) 19 24.
3. Ashley McGillivray, 'Deepfake and Human Rights: The Non Consensual Synthesis of Intimate Images' (2020) *International Journal of Law and Information Technology*, 28(3), 225, 230.
4. Danielle Keats Citron, *Hate Crimes in Cyberspace* (Harvard University Press, 2014) 35.
5. Section 78 of the Bharatiya Nyaya Sanhita, 2023 now defines stalking; previously, the Indian Penal Code, 1860 contained s 354D.
6. Diana Freed et al., "A Stalker's Paradise": How Intimate Partner Abusers Exploit Technology (2018) *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* 1, 46.
7. Nicolas Suzor et al., (2019) *Human Rights by Design: The Responsibilities of Social Media Platforms to Address Gender Based Violence Online* *Policy & Internet*, 11(1), 84, 88.
8. The term "AI social media nexus" is used here to denote the reciprocal amplification between algorithmic systems and platform architectures; see Shoshana Zuboff, *The Age of Surveillance Capitalism* (PublicAffairs, 2019) 270 275.
9. Freed et al. (n 6) 5.
10. Ibid 7 8.
11. Zuboff (n 8) 93.
12. For a description of the data broker market in India, see Privacy International, *State of Privacy India* (2019) 12 14.
13. Citron (n 4) 51 52.
14. Suzor et al. (n 7) 94.
15. Amnesty International (n 2) 30 34.
16. Freed et al. (n 6) 56; predictive analytics in stalking is an emerging field, discussed extensively in Anastasia Powell and Nicola Henry, *Sexual Violence in a Digital Age* (Palgrave Macmillan, 2017) 58 62.
17. Powell and Henry (n 17) 59.
18. McGillivray (n 3) 228.
19. Id at 232.
20. Nicola Henry and Anastasia Powell, 'Technology Facilitated Sexual Violence: A Literature Review of Empirical Research' (2018) *Trauma, Violence, & Abuse* 19(2), 195, 201.
21. Alice E. Marwick and Robyn Caplan, 'Drinking Male Tears: Language, the Manosphere, and Networked Harassment' (2018) *Feminist Media Studies* 18(4), 543, 548 554.
22. Ibid.
23. Amnesty International (n 2) 34.
24. Puttaswamy (n 1) para 298 (Chandrachud J).

25. Id, para 325.
26. Id, para 303.
27. See Francis Coralie Mullin v. Administrator, Union Territory of Delhi (1981) 1 SCC 608, para 3; Puttaswamy (n 1) para 144.
28. UN Committee on the Elimination of Discrimination against Women, General Recommendation No. 35 on Gender Based Violence against Women, UN Doc CEDAW/C/GC/35 (2017) para 20.
29. McGillivray (n 3) 245.
30. CEDAW Committee (n 34) para 31.
31. Human Rights Council, Report of the Special Rapporteur on Violence against Women, its Causes and Consequences on Online Violence against Women and Girls from a Human Rights Perspective, UN Doc A/HRC/38/47 (2018) para 29.
32. Shreya Singhal v. Union of India (2015) 5 SCC 1, para 21.
33. For the horizontal application of fundamental rights, see Puttaswamy (n 1) para 109.
34. This intersectional pattern is documented in Amnesty International (n 2) 44 45.
35. Dalit women's specific vulnerability is highlighted in various NGO reports; see e.g., National Dalit Movement for Justice, Caste and Cyber Violence (2020) 8 10.
36. Council of Europe, Cyberviolence against Women and Girls: A World Wide Wake Up Call (Council of Europe Publishing, 2021) 12.
37. The Information Technology Act, 2000 (Act No. 21 of 2000), ss 66E, 67, 67A.
38. For enforcement challenges, see Centre for Internet and Society, Cyber Stalking in India: A Report (2021) 22 28.
39. IT Rules, 2021, r 3(2)(b), (d).
40. Bharatiya Nyaya Sanhita, 2023, s 78.
41. Section 78(1) defines stalking as "following a person and contacting, or attempting to contact such person to foster personal interaction repeatedly despite a clear indication of disinterest."
42. Sections 294 296 deal with obscenity, word, gesture or act intended to insult modesty, and voyeurism, respectively.
43. Digital Personal Data Protection Act, 2023, ch II, ss 6 11.
44. Section 6 of the DPDP Act requires consent for processing, but there is no specific prohibition on processing personal data for stalking purposes.
45. The Schedule to the DPDP Act lists "psychological harm" and "loss of reputation" as harms but leaves the enforcement mechanism to the Board.
46. The concept of a constitutional tort for violation of Article 21 was recognised in Nilabati Behera v. State of Orissa (1993) 2 SCC 746.
47. Puttaswamy (n 1) para 325.
48. Puttaswamy (n 1) para 325; Human Rights Council (n 1) para 67.
