INTERNATIONAL JOURNAL FOR LEGAL RESEARCH AND ANALYSIS

Open Access, Refereed Journal Multi Disciplinary
Peer Reviewed

www.ijlra.com

ISSN: 2582-6433

DISCLAIMER

The views, opinions, interpretations, and conclusions expressed in the articles published in this journal are solely those of the respective authors. They do not necessarily reflect the views of the Editorial Board, Editors, Reviewers, Advisors, or the Publisher of IJLRA.

Although every reasonable effort has been made to ensure the accuracy, authenticity, and proper citation of the content published in this journal, neither the Editorial Board nor IJLRA shall be held liable or responsible, in any manner whatsoever, for any loss, damage, or consequence arising from the use, reliance upon, or interpretation of the information contained in this publication.

The content published herein is intended solely for academic and informational purposes and shall not be construed as legal advice or professional opinion.

ISSN: 2582-6433

ABOUT US

The *International Journal for Legal Research & Analysis (IJLRA)* (ISSN: 2582-6433) is a peer-reviewed, academic, online journal published on a monthly basis. The journal aims to provide a comprehensive and interactive platform for the publication of original and high-quality legal research.

IJLRA publishes Short Articles, Long Articles, Research Papers, Case Comments, Book Reviews, Essays, and interdisciplinary studies in the field of law and allied disciplines. The journal seeks to promote critical analysis and informed discourse on contemporary legal, social, and policy issues.

The primary objective of IJLRA is to enhance academic engagement and scholarly dialogue among law students, researchers, academicians, legal professionals, and members of the Bar and Bench. The journal endeavours to establish itself as a credible and widely cited academic publication through the publication of original, well-researched, and analytically sound contributions.

IJLRA welcomes submissions from all branches of law, provided the work is original, unpublished, and submitted in accordance with the prescribed submission guidelines. All manuscripts are subject to a rigorous peer-review process to ensure academic quality, originality, and relevance.

Through its publications, the *International Journal for Legal Research & Analysis* aspires to contribute meaningfully to legal scholarship and the development of law as an instrument of justice and social progress.

ISSN: 2582-6433

PUBLICATION ETHICS, COPYRIGHT & AUTHOR RESPONSIBILITY STATEMENT

The *International Journal for Legal Research and Analysis (IJLRA)* is committed to upholding the highest standards of publication ethics and academic integrity. All manuscripts submitted to the journal must be original, unpublished, and free from plagiarism, data fabrication, falsification, or any form of unethical research or publication practice. Authors are solely responsible for the accuracy, originality, legality, and ethical compliance of their work and must ensure that all sources are properly cited and that necessary permissions for any third-party copyrighted material have been duly obtained prior to submission. Copyright in all published articles vests with IJLRA, unless otherwise expressly stated, and authors grant the journal the irrevocable right to publish, reproduce, distribute, and archive their work in print and electronic formats. The views and opinions expressed in the articles are those of the authors alone and do not reflect the views of the Editors, Editorial Board, Reviewers, or Publisher. IJLRA shall not be liable for any loss, damage, claim, or legal consequence arising from the use, reliance upon, or interpretation of the content published. By submitting a manuscript, the author(s) agree to fully indemnify and hold harmless the journal, its Editor-in-Chief, Editors, Editorial Board, Reviewers, Advisors, Publisher, and Management against any claims, liabilities, or legal proceedings arising out of plagiarism, copyright infringement, defamation, breach of confidentiality, or violation of third-party rights. The journal reserves the absolute right to reject, withdraw, retract, or remove any manuscript or published article in case of ethical or legal violations, without incurring any liability.

ISSN: 2582-6433

DEEPFAKES AS DIGITAL DEVIANCE: A CRIMINOLOGICAL ANALYSIS OF OFFENDER MOTIVATION

AUTHORED BY - PRANAV GOEL & ABHINAV SARASWAT

ABSTRACT

The rapid rise of deepfake technology poses a significant and escalating threat to our digital society. By using artificial intelligence to create highly realistic yet entirely fake videos and audio, this technology has opened the door to new and disturbing forms of crime. It can be used to spread political misinformation, commit sophisticated financial fraud, harass individuals and create non-consensual pornography, causing profound emotional and reputational harm. While existing literature has predominantly addressed the victimological consequences, such as reputational harm and privacy violation, there remains a critical gap in the criminological tendencies of those who create and disseminate such content. From a legal standpoint, deepfakes test the adequacy of existing cybercrime frameworks, while from a criminological perspective, they provide a ground to examine the motivations and behaviours of modern digital offenders. This project seeks to bridge this gap by analysing deepfake creation within established criminological theory.

The research project has two central questions: (1) How can classical, psychological, and sociological criminological theories explain the behavioural tendencies and motivations of deepfake creators? (2) To what extent does deepfake creation exemplify a form of digital deviance that is rationalised through techniques of neutralisation and reinforced by online subcultural association? This project hypothesises that the act of creating a malicious deepfake is a deliberate choice, influenced by a combination of individual psychology, social environment, and rational calculation.

The project is structured into four substantive chapters. The first chapter introduces deepfake technology and its treatment under Indian and comparative legal frameworks, highlighting legal and enforcement gaps that necessitate a criminological analysis. The second chapter examines individualistic approaches to deepfake offences, drawing on classical rational choice theory, routine activity theory and psychological perspectives. It also considers illustrative

ISSN: 2582-6433

case studies of offender motives: the Rana Ayyub deepfake pornography case in India, which highlighted targeted harassment rooted in misogyny; the Gabriel Krüger case in Germany, where financial fraud was the central driver; and U.S. election-related deepfake scandals, where political misinformation was weaponised to achieve ideological influence. The third chapter elaborates on deepfake creation within sociological frameworks, engaging with differential association, neutralisation, and subcultural theories to explain the social processes that sustain this conduct online. The fourth chapter synthesises these approaches to propose a criminological offender typology, categorising deepfake creators as occasional, professional, or chronic offenders, and considers broader implications for criminal law and policy.

By mapping deepfake offending onto established theories and identifying patterns of offender rationalisation, the project aims to provide a more holistic understanding of this digital deviance. It seeks to underscore that the phenomenon of deepfakes is not merely a technological anomaly but a criminological reality that requires robust theoretical and legal engagement.

Keywords: Deepfakes, Criminological Theory, Digital Deviance, Neutralisation, Cybercrime.

INTRODUCTION

The emergence of deepfake technology has transformed the digital landscape, enabling the creation of highly realistic yet entirely fabricated videos and audio. While initially developed for good purposes such as entertainment and education, this technology has increasingly been exploited to commit crimes, including political misinformation, financial fraud, targeted harassment, and non-consensual pornography.[1] These activities result in profound social, emotional, and reputational consequences for victims.[2]

Existing research has largely focused on the victimological impact of deepfakes, such as privacy violations and reputational harm. However, there remains a significant gap in understanding the criminological dimension, specifically, the motivations, behaviours, and decision-making processes of those who create and disseminate deepfakes. Examining these

[1] Fatih Arslan, 'Deepfake Technology: A Criminological Literature Review' (2023) 11(1) Sakarya Journal of Law 701, 702.
[2] Dhyey Sadiwala, 'The Case for Criminalizing Deepfake AI' (2025) 10(1) International Journal of Novel Research and Development 729, 732.

ISSN: 2582-6433

actions through the lens of established criminological theories is essential to conceptualising deepfake creation as a form of digital deviance rather than merely a technological anomaly.[3]

To address this gap, the project is structured into four substantive chapters. The first chapter introduces deepfake technology and its treatment under Indian and comparative legal frameworks, highlighting regulatory and enforcement gaps that necessitate a criminological analysis. The second chapter examines individual-level approaches to deepfake offences, drawing on classical rational choice theory, routine activity theory, and psychological perspectives. It also uses illustrative case studies, including the Rana Ayyub deepfake pornography case, the Gabriel Krüger financial fraud case, and U.S. election-related deepfake incidents, to demonstrate offender motivations.[4]

The third chapter engages with sociological explanations of deepfake creation, utilising differential association, neutralisation, and subcultural theories to explain the social processes that sustain deviant online conduct. The fourth chapter synthesises these approaches to propose a criminological typology of deepfake offenders, categorising them as occasional, professional, or chronic, while also considering broader implications for criminal law and policy. This project, therefore, seeks to bridge the existing theoretical and empirical gaps by analysing deepfake creation through a criminological lens.

I. DEEPFAKE TECHNOLOGY AND LEGAL FRAMEWORKS

The proliferation of deepfake technology represents a watershed moment in the evolution of digital crime. Deepfakes, created through generative adversarial networks ("**GANs**") and other machine learning algorithms, enable the production of synthetic media that is increasingly indistinguishable from authentic content.[5] This technological capability has evolved rapidly since 2017, when the term "deepfake" first emerged on social media platforms.[6] The accessibility of deepfake creation tools has democratised a form of manipulation previously

[3] Julia Stavola and Kyung-Shick Choi, 'Victimization by Deepfake in the Metaverse: Building a Practical Management Framework' (2023) 6(2) International Journal of Cybersecurity Intelligence & Cybercrime; BNP Panda, 'Deepfake Technology in India and World: Foreboding and Forbidding' (2025) Legal Horizons Quarterly Review.
[4] Ibid.
[5] Bobby Chesney and Danielle Citron, "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security" (2019) 107(1) California Law Review 1760; Peipeng Yu et al., "A Survey on Deepfake Video Detection," *IET Biometrics* 10/6 (November 2021), 607-624.
[6] Matthew B. Kugler - Carly Pace, "Deepfake Privacy: Attitudes and Regulation," *SSRN Electronic Journal* 116/3 (2021), 611-680.

ISSN: 2582-6433

requiring substantial technical expertise and resources, thereby expanding the pool of potential offenders.[7]

A Understanding Deepfake Technology

Deepfake technology operates through sophisticated artificial intelligence systems that analyse and replicate human facial features, voice patterns, and bodily movements. The technical process involves training neural networks on extensive datasets of images, videos, or audio recordings of a target individual, subsequently generating new content that convincingly portrays that person engaging in fabricated activities or making statements they never uttered. Recent advancements have reduced the barrier to entry, with user-friendly applications and open-source software enabling even non-technical users to create convincing deepfakes within hours.[8]

The malicious applications of this technology are diverse and deeply concerning. Non-consensual intimate imagery, particularly targeting women, constitutes a significant proportion of deepfake content circulating online.[9] Research indicates that approximately 96% of deepfake videos available on the internet are pornographic in nature, with the overwhelming majority featuring women without their consent.[10] The high quality image generation in chatbots' integration in social media platform itself allows more easy accessibility.[11] Beyond sexual exploitation, deepfakes have been weaponised for political manipulation, financial fraud through impersonation, and reputational destruction of public figures and private individuals alike.[12]

B Indian Legal Framework

India's legal and regulatory framework has evolved to address deepfake technology, though gaps remain. The Digital Personal Data Protection Act 2023 ("**DPDPA**") mandates consent for

[7] BNP Panda, 'Deepfake Technology in India and World: Foreboding and Forbidding' (2025) Legal Horizons Quarterly Review.
[8] Fatih Arslan, 'Deepfake Technology: A Criminological Literature Review' (2023) 11(1) Sakarya Journal of Law 701, 703.
[9] Alex Schuler, *Deep Fake* (Atlantic Books 2025).
[10] Deepfake Statistics 2025: AI Fraud Data & Trends (DeepStrike, 2025) https://deepstrike.io/blog/deepfake-statistics-2025 accessed 01 October 2025.
[11] Amelia Gentleman and Helena Horton, '"Add Blood, Forced Smile": How Grok's Nudification Tool Went Viral' *The Guardian* (11 January 2026); Kate Conger, Lizzie DeardenKate Conger reported from San Francisco and Lizzie Dearden from London, 'Elon Musk's A.I. Is Generating Sexualized Images of Real People, Fueling Outrage' *The New York Times* (9 January 2026).
[12] Deepfakes | Current Affairs (VisionIAS, 2025) https://visionias.in/current-affairs/monthly-magazine/2025-06-17/science-and-technology/deepfakes accessed 03 October 2025.

ISSN: 2582-6433

personal data processing under Section 6,[13] classifying non-consensual deepfake use as a breach subject to fines of up to ₹250 crore, complementing Section 79 of the Information Technology Act ("**IT Act**")[14] through data minimisation and fiduciary duties. The Bharatiya Nyaya Sanhita 2023 ("**BNS**"), modernises offences: Section 353 criminalises statements conducing to public mischief, Section 111 targets organised cybercrimes, Sections 319 and 336 address personation and electronic forgery, and Section 356 extends defamation provisions to synthetic media.[15] Existing IT Act provisions, particularly Section 67 for obscene material,[16] continue to apply, while the Bharatiya Sakshya Adhiniyam 2023 ("**BSA**") Section 63 requires authentication of electronic records, including hash and source verification, critical for admissibility of deepfake evidence.[17]

Regulatory efforts supplement these statutes. In November 2023, the Ministry of Electronics and Information Technology ("**MeitY"**) advised platforms to remove deepfakes within 36 hours, targeting gendered harms,[18] while the March 2024 guidelines mandated persistent labels and metadata to enable originator tracing.[19] CERT-In in November 2024 recommended AI/ML detection tools, watermarking, source verification, and adoption of C2PA standards, with trials ongoing in 2025.[20] The Election Commission of India ("**ECI**"), in May 2024, directed political parties to remove deepfake posts within three hours during the Model Code of Conduct period to prevent AI-driven misinformation.[21] Most significantly, MeitY notified the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026 on February 10, 2026 (effective February 20, 2026),[22] constituting the first formal statutory recognition of synthetically generated information ("SGI") under Indian law. The 2026 Amendment Rules require all intermediaries offering AI content generation tools to deploy technical measures preventing unlawful synthetic content, including non-consensual intimate imagery ("NCII") and child sexual abuse material, mandatory prominent labelling and permanent metadata embedding for all public-facing SGI and imposes enhanced due diligence obligations on Significant Social Media Intermediaries ("SSMIs"), including the requirement

[13] *Digital Personal Data Protection Act 2023*, s 6.
[14] *Information Technology Act 2000*, s 79.
[15] *Bharatiya Nyaya Sanhita 2023*, ss 111, 319, 336, 353, and 356.
[16] *Information Technology Act 2000*, s 67.
[17] *Bharatiya Sakshya Adhiniyam 2023*, s 63.
[18] Ministry of Electronics and Information Technology, *Advisory on Deepfakes* (November 2023).
[19] Ministry of Electronics and Information Technology, *Guidelines on Labeling and Metadata for Synthetic Media* (March 2024).
[20] Indian Computer Emergency Response Team (CERT-In), *Recommendations on Deepfake Detection and C2PA Standards* (November 2024).
[21] Election Commission of India, *Directive to Political Parties on Removal of Deepfake Posts during the Model Code of Conduct Period* (May 2024).
[22] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules 2026.

ISSN: 2582-6433

to obtain user declarations for uploaded content. Critically, the general content takedown timeline has been tightened from 36 hours to three hours, and safe harbour protection under Section 79 of the IT Act is now expressly conditioned upon compliance with these obligations, making proactive compliance a core risk management function for platforms.

Notwithstanding the landmark 2026 Amendment Rules, India's approach continues to operate across multiple instruments, with no single consolidated statute explicitly addressing synthetic media manipulation. The IT Act remains foundational for privacy violations under Section 79 and obscene content under Section 67. However, having predated deepfake technology, the Act itself still lacks intrinsic specificity, a deficit now partially remedied through the 2026 Amendment Rules to the IT Rules 2021, which supply the technology-specific obligations that the bare Act could not. The BNS recognises sexual deepfakes but largely ignores political disinformation, fraud, and other non-sexual uses.[23] The 2026 Amendment Rules represent a material shift from reactive, advisory-based enforcement toward mandatory, proactive compliance and making platforms now legally obligated to deploy automated detection tools, enforce mandatory labelling, and meet enforceable takedown timelines, with safe harbour loss as the operative sanction for non-compliance. Nevertheless, significant gaps persist including non-sexual deepfake harms such as political disinformation and financial fraud remain outside dedicated statutory coverage, and the BNS provisions addressing these uses are still limited in scope. Overall, the evolving legislative and regulatory landscape reflects meaningful progress alongside the continuing need for a comprehensive, technology-specific statute that consolidates and expands upon the piecemeal framework currently in force.[24]

C Comparative Legal Perspectives

International jurisdictions have adopted various approaches to regulating deepfakes, which provide helpful guidance for the Indian legislators. The European Union's Digital Services Act of 2022 establishes comprehensive obligations for digital platforms regarding content moderation, transparency, and risk assessment, with specific provisions applicable to synthetic media.[25] The Act mandates the clear labelling of manipulated content and imposes significant penalties for non-compliance, reflecting a preventive rather than purely punitive approach.

In the United States, legislative responses have primarily emerged at the state level.

[23] India well-equipped to tackle evolving online harms and deepfakes (Press Information Bureau, 2023); Legal Dimensions of Deepfake Technology: Privacy, Consent, and Criminal Liability (Juris Centre, 2025).
[24] *Aditya Mehrotra*, Dissecting the Framework of Deep Fakes in India; Deepfakes And The Law: The Need Of The Hour Is An Effective Legal Framework In India (Mondaq, 2025).
[25] European Union, Digital Services Act 2022.

ISSN: 2582-6433

California's Assembly Bill 730, enacted in 2019, criminalises the creation and distribution of deepfakes related to elections within 60 days of voting. Similarly, Texas and Virginia have introduced legislation specifically addressing deepfake pornography, recognising it as a distinct category of digital sexual abuse. At the federal level, the DEEPFAKES Accountability Act and the Malicious Deep Fake Prohibition Act have been proposed but have not yet been enacted, indicating ongoing legislative deliberation.[26]

China has implemented notably stringent regulations through its Deep Synthesis Provisions, which took effect in January 2023.[27] These regulations require deepfake content to be clearly labelled and mandate platforms to implement technical measures for detection and verification. Additionally, they establish criminal liability for creators and distributors of malicious deepfakes, supported by China's robust cybersecurity enforcement framework.

Therefore, it can be said that effective regulation of deepfakes requires a multi-faceted approach encompassing criminal sanctions, platform accountability, technological safeguards, and public awareness. India's current regulatory framework falls short of this comprehensive model, particularly in the areas of proactive detection, platform liability, and addressing non-sexual deepfake offences.

II. INDIVIDUAL-LEVEL CRIMINOLOGICAL EXPLANATIONS

Understanding why individuals create malicious deepfakes necessitates examination through classical and psychological criminological theories that focus on individual decision-making, opportunity structures, and cognitive processes.[28]

A Classical Rational Choice Theory

Classical criminology, which originated with Cesare Beccaria and Jeremy Bentham, suggests that criminal behaviour arises from rational calculations in which individuals weigh the potential benefits against the risks of detection and punishment.[29] When applied to the creation of deepfakes, this framework indicates that offenders conduct a cost-benefit analysis before producing and distributing synthetic media.[30]

The perceived benefits of creating deepfakes differ based on the offender's motivations. For

[26] California Assembly Bill 730 (2019); DEEPFAKES Accountability Act (US Congress, proposed); Malicious Deep Fake Prohibition Act (US Congress, proposed).
[27] Deep Synthesis Provisions (China, 2023).
[28] Fatih Arslan, 'Deepfake Technology: A Criminological Literature Review' (2023) 11(1) Sakarya Journal of Law 701, 703.
[29] Ahmad Siddique, *Criminology, Penology and Victimology* (8th edn, Eastern Book Company 2024).
[30] H Etienne, 'The future of online trust (and why Deepfake is advancing it)' (2021).

ISSN: 2582-6433

individuals engaged in sexual exploitation, gratification may come from humiliating victims, exerting power, or fulfilling voyeuristic desires. Political actors might determine that the reputational damage they inflict on opponents or the viral spread of disinformation justifies the risks of being exposed. Financial fraudsters evaluate potential monetary gains against the likelihood of detection and prosecution.[31]

Importantly, the risk aspect of this calculation often favours offenders for several reasons. Anonymity provided by internet platforms and encryption technologies decreases the chances of identification. Additionally, the early stage of deepfake detection technology and the limited technical capabilities of law enforcement agencies further lessen perceived risks. The cross-jurisdictional nature of online offences complicates prosecution, especially when offenders operate from regions with weak cybercrime enforcement.[32]

Where sanctions do exist, they may be insufficient to deter determined offenders, particularly when substantial perceived benefits are at stake. The Gabriel Krüger case in Germany illustrates this rational calculation; a deepfake was used to impersonate a corporate executive, resulting in the defrauding of a company of millions. Despite eventual detection and prosecution, the initial success of the scheme and the significant amount obtained show how rational actors might view deepfake fraud as a worthwhile pursuit.[33]

B Routine Activity Theory

Routine Activity Theory, developed by Lawrence Cohen and Marcus Felson, suggests that crime occurs when three elements come together: a motivated offender, a suitable target, and the absence of capable guardianship.[34] This framework is particularly useful for understanding deepfake offences within the digital landscape.

Motivated offenders are plentiful in online spaces, driven by various motivations, including sexual gratification, financial gain, political ideology, or simple malice. The availability of deepfake creation tools has significantly widened the pool of potential offenders beyond those with advanced technical skills. Open-source software, mobile applications, and online tutorials have made deepfake production more accessible, lowering the barriers to entry.

Similarly, suitable targets are abundant. Public figures, celebrities, and politicians are especially vulnerable due to the vast amount of their images and videos available online, which

[31] H Bashir, 'Rational Choices, Moral Emotions, and Social Media Literacy' (2024).
[32] M Appel, 'The detection of political deepfakes' (2022) 27(4) Journal of Computer-Mediated Communication.
[33] BNP Panda, 'Deepfake Technology in India and World: Foreboding and Forbidding' (2025) Legal Horizons Quarterly Review.
[34] Peter Joyce, *Criminology: A Complete Introduction: Teach Yourself* (Teach Yourself 2018).

ISSN: 2582-6433

serve as training data for deepfake algorithms. However, private individuals are increasingly being targeted as well, particularly in cases of intimate partner violence, workplace harassment, or personal vendettas. The Rana Ayyub case illustrates how a high-profile journalist was targeted for misogynistic harassment through deepfake pornography, where her public visibility and controversial reporting made her susceptible to such attacks.[35]

The absence of capable guardianship is perhaps the most crucial factor enabling deepfake offences. Digital platforms, despite having content moderation capabilities, often struggle to detect and remove deepfakes promptly. Detection technologies lag behind creation capabilities, and the sheer volume of content uploaded daily surpasses the capacity for thorough human review. Law enforcement agencies also typically lack the technical expertise and resources necessary for the effective investigation and prosecution of deepfake offences.[36]

Moreover, the social guardianship usually provided by bystanders and communities in physical spaces is weakened online. The anonymity and distance of digital interactions reduce informal social controls, and some online communities may even promote deviant behaviour rather than discourage it. This is particularly evident in forums and platforms where deepfake pornography is shared, creating echo chambers that normalise and celebrate such violations.

C Psychological Criminology

Psychological criminology explores the individual traits, cognitive processes, and mental states that contribute to criminal behaviour.[37] Understanding the psychological factors involved in creating deepfakes is important.

One relevant concept is moral disengagement, developed by Albert Bandura. This refers to the cognitive mechanisms that allow individuals to justify harmful actions to themselves.[38] Deepfake creators may use various strategies for moral disengagement, such as dehumanising their victims, shifting responsibility onto technology or others, minimising the harm caused, or appealing to higher loyalties like political causes or group solidarity.[39]

The online disinhibition effect also plays a significant role. In digital environments, individuals often behave with less restraint than they would in face-to-face interactions. Anonymity, invisibility, and asynchronicity reduce accountability and empathy, which enables individuals

[35] Legal Dimensions of Deepfake Technology: Privacy, Consent, and Criminal Liability (Juris Centre, 2025).
[36] *Aditya Mehrotra*, Dissecting the Framework of Deep Fakes in India.
[37] Ahmad Siddique, *Criminology, Penology and Victimology* (8th edn, Eastern Book Company 2024).
[38] N.V. Paranjape, *Criminology & Penology Including Victimology* (19th edn, Eastern Book Company 2023)
[39] Albert Bandura, 'Selective Moral Disengagement in the Exercise of Moral Agency' (2002) 80(2) Journal of Moral Education 101.

ISSN: 2582-6433

to produce content they might avoid if they faced immediate social consequences.[40]

Certain personality traits are linked to online offending behaviours. Research shows that individuals who exhibit high levels of Dark Triad traits, narcissism, Machiavellianism, and psychopathy are more likely to engage in cyberbullying, harassment, and other forms of digital aggression. These traits may also predispose individuals to create deepfakes, particularly in scenarios related to revenge pornography or targeted harassment.

Gender-based attitudes, especially hostile sexism and misogyny, are significant factors in the creation of deepfake pornography. The overwhelming targeting of women in non-consensual intimate deepfakes reflects broader patterns of technology-facilitated gender-based violence. Psychological research on sexual aggression and objectification provides frameworks for understanding how some individuals come to see women as appropriate targets for such violations.

D Case Study Analysis

The following cases highlight the diverse motivations and contexts of deepfake offences, demonstrating the relevance of individual-level criminological theories.

The first case involves Rana Ayyub, an Indian journalist who became the target of deepfake pornography after her critical reporting on politically sensitive issues. This case exemplifies how deepfakes can be used as tools for gendered harassment and political intimidation. Offenders, including political opponents and misogynists, calculated that the reputational harm and psychological distress inflicted upon Ayyub would silence or punish her. The presence of a suitable target (a visible female journalist), the motivation of the offenders, and the lack of effective guardianship (platform inaction and impunity) converged to enable this violation.

The second case is that of Gabriel Krüger, which illustrates the use of deepfakes in corporate crime. In this scenario, voice synthesis technology was employed to impersonate a company executive, instructing an employee to transfer a substantial sum of money to fraudulent accounts. The offender's rational calculation centred on financial gain, with the sophistication of the deepfake technology providing confidence in achieving success. This case shows how deepfakes expand the toolkit of white-collar criminals by taking advantage of trust relationships and institutional hierarchies.

Lastly, incidents of deepfakes related to U.S. elections, including fabricated videos of political candidates, reveal ideological motivations. Offenders in these cases aim to influence electoral

[40] John Suler, 'The Online Disinhibition Effect' (2004) 7(3) CyberPsychology & Behavior 321.

ISSN: 2582-6433

outcomes, damage opponents' reputations, or create general distrust in democratic processes. Their calculations involve weighing the potential political impact against the likelihood of detection before the content's intended purpose is served. Often, even brief viral circulation can achieve the desired effect before the content is removed or debunked.[41]

III. SOCIOLOGICAL EXPLANATIONS OF DEEPFAKE CREATION

While individual-level theories illuminate personal motivations and calculations, sociological perspectives are essential for understanding how social structures, group processes, and cultural contexts shape and sustain deepfake offending.[42]

A Differential Association Theory

Edwin Sutherland's Differential Association Theory suggests that criminal behaviour is learned through interactions with others, particularly within intimate personal groups.[43] Individuals not only learn the techniques of committing crimes but also the motives, rationalisations, and attitudes that favour violating the law.[44]

Online communities focused on deepfake creation function as virtual schools for this type of crime, where learning processes take place. Forums, Discord servers, reddit communities and other platforms facilitate discussions about technical methods, share software tools, and circulate created content. New members are introduced to the norms and values of these communities, which often portray deepfake creation as harmless entertainment, an opportunity for technological experimentation, or even justified revenge.

The intensity, duration, and priority of these associations affect how strongly individuals internalise pro-deepfake attitudes. Regular involvement in these communities, along with exposure to rationalising narratives and positive reinforcement for creating and sharing deepfakes, reinforces commitment to this deviant behaviour. Additionally, the anonymity and global reach of online platforms allow individuals to connect with like-minded peers regardless of their geographic location, further amplifying the influence of deviant peer groups.[45]

[41] M Pawelec, 'Deepfakes and Democracy (Theory): How Synthetic Audio-Visual Media Distort Deliberation, Trust, and Democratic Legitimacy' (2022).
[42] Fatih Arslan, 'Deepfake Technology: A Criminological Literature Review' (2023) 11(1) Sakarya Journal of Law 701, 715.
[43] Ahmad Siddique, *Criminology, Penology and Victimology* (8th edn, Eastern Book Company 2024).
[44] A Eberl, 'Using deepfakes for experiments in the social sciences' (2022) Frontiers in Sociology.
[45] QJ Ullrich, 'Is This Video Real? The Principal Mischief of Deepfakes and How the Lanham Act Can Address It' (2022) 55 Columbia Journal of Law & Social Problems 1, 4.

ISSN: 2582-6433

B Techniques of Neutralisation

Gresham Sykes and David Matza's Neutralisation Theory explain how offenders use cognitive techniques to rationalise deviant behaviour, temporarily suspending their moral constraints.[46] Deepfake creators frequently employ five primary techniques of neutralisation[47]:

1. **Denial of Responsibility:** Offenders attribute blame to technology itself ("the AI made it easy") or to the victims ("she shouldn't have put photos online"). This deflection obscures the intentional choices involved in creating and distributing harmful content.
2. **Denial of Injury:** They minimise the harm caused by asserting, "It's just pixels," "No one actually got hurt," or "It's obviously fake, so it doesn't matter." Such rationalisations ignore the profound psychological, reputational, and material harm that deepfakes inflict on victims.
3. **Denial of the Victim:** This involves victim-blaming and dehumanisation. Public figures may be portrayed as deserving targets for deepfake attacks because of their fame, wealth, or controversial opinions. Women targeted with deepfake pornography are often blamed for their appearance, behaviour, or online presence, transforming victims into perceived legitimate targets.
4. **Condemnation of the Condemners:** Offenders dismiss those who criticise deepfake creation as censorious, humourless, or hypocritical. They might argue that other forms of image manipulation or misinformation exist, claiming that deepfakes are unfairly singled out.
5. **Appeal to Higher Loyalties:** This technique justifies deepfakes through references to group solidarity, political causes, or principles of free expression. Political deepfakes are rationalised as necessary for exposing truths or combating perceived threats. Within close-knit groups, creating deepfakes may be framed as an act of loyalty to friends or community norms.

C Subcultural Theory

Subcultural criminology explores how distinct value systems and behavioural norms within subgroups facilitate and legitimise deviant behaviour. Online spaces dedicated to deepfake creation have developed into subcultures with their own status hierarchies, jargon, rituals, and

[46] N.V. Paranjape, *Criminology & Penology Including Victimology* (19th edn, Eastern Book Company 2023).
[47] Akanksha Wadhawan, 'Deepfake Technology as a New Tool for Criminal Offences: Legal Challenges and Its Way Forward in Criminal Law' (2025) 7(3) International Journal for Multidisciplinary Research.

ISSN: 2582-6433

normative structures.[48]

These subcultures often arise from broader internet communities characterised by transgressive humour, anti-establishment attitudes, and hostility toward feminist and progressive movements. Within these spaces, deepfake creation is celebrated as a technical skill, a form of comedic creativity, or a type of political activism rather than being viewed as criminal behaviour.

Status within deepfake subcultures is awarded to those who demonstrate technical proficiency, create particularly convincing or shocking content, or successfully evade detection and consequences. This competition for status drives ongoing engagement with deepfake creation, as individuals seek recognition and prestige within their communities.

The values of these subcultures often include disdain for mainstream norms, a celebration of transgression, and hostility toward victims and authorities. These oppositional values insulate members from external criticism and strengthen internal cohesion, making intervention and desistance more challenging.

The intersection of misogynistic subcultures and deepfake technology is especially concerning. Communities built around "manosphere" ideologies, incel belief systems, or general anti-feminism have embraced deepfake pornography as a means of punishing, controlling, and objectifying women. Within these subcultures, the creation and sharing of non-consensual intimate deepfakes are framed not as abuse, but as justified resistance against feminism or as entertainment.

IV. Toward a Criminological Typology of Deepfake Offenders

Synthesising the individual and sociological perspectives examined above enables the development of a criminological typology that categorises deepfake offenders according to their behavioural patterns, motivations, and degree of commitment to offending.[49]

A Occasional Offenders

Occasional offenders create deepfakes infrequently, typically in response to specific situations or emotional states. These individuals usually do not have a consistent commitment to producing deepfakes and may only make one or a few. Their motivations are often personal

[48] Deepfake Statistics (Eftsure, 2025) https://www.eftsure.com/statistics/deepfake-statistics/ accessed 0963 October 2025.

[49] Fatih Arslan, 'Deepfake Technology: A Criminological Literature Review' (2023) 11(1) Sakarya Journal of Law 701, 715.

ISSN: 2582-6433

and reactive, such as seeking revenge after a relationship ends, harassing specific individuals, or participating impulsively in online "pranks."[50]

From a rational choice perspective, occasional offenders may not engage in a thorough cost-benefit analysis; instead, they act based on immediate emotions or opportunities. Their ability to create deepfakes is often made easier by user-friendly applications that require little technical skill, which lowers the barrier for impulsive behaviour. These offenders are less likely to be part of deviant subcultures but may temporarily use neutralisation techniques to justify their actions.[51]

Occasional offenders present promising opportunities for deterrence and intervention. They tend to maintain connections to conventional values and are more likely to feel shame and social disapproval. Legal consequences, educational programs that emphasise the harm caused by their actions, and technological barriers could effectively prevent or stop their offending behaviour.[52]

B Professional Offenders

Professional offenders view the creation of deepfakes as a strategic activity aimed at achieving specific goals, often financial in nature. This group includes individuals who produce deepfakes for hire, engage in fraud or extortion, or create commercial deepfake pornography. Their activities are calculated and persistent, characterised by advanced technical skills and a strong emphasis on operational security.

Rational choice theory helps to explain the behaviour of these professional offenders, who conduct thorough cost-benefit analyses. They determine that engaging in deepfake activities presents an acceptable risk-reward balance. As a result, they invest in technologies and techniques that improve quality while minimising the chances of detection. Often, these offenders operate across multiple jurisdictions to take advantage of legal loopholes and limitations in law enforcement.[53]

Unlike other offenders, professional deepfake creators are less influenced by subcultural values and are more motivated by economic incentives and professional networks. They typically maintain an emotional distance from their victims, approaching deepfake creation as a business

[50] Thomas Brewster, 'Fraudsters Cloned Company Director's Voice in $35 Million Bank Heist, Police Find' Forbes (2021).
[51] Akanksha Wadhawan, 'Deepfake Technology as a New Tool for Criminal Offenses: Legal Challenges and Its Way Forward in Criminal Law' (2025) 7(3) International Journal for Multidisciplinary Research 1, 5.
[52] Clementina Salvi, 'Deepfake Evidence in Criminal Proceedings' (2024) AI and Criminal Justice.
[53] Deepfake Statistics & Trends 2025 | Key Data & Insights (Keepnet Labs, 2025) https://keepnetlabs.com/blog/deepfake-statistics-and-trends accessed 05 October 2025.

ISSN: 2582-6433

rather than a form of personal expression or ideological commitment. This focus on instrumental goals makes them responsive to measures aimed at reducing opportunities for offending and increasing punitive actions. However, their sophisticated methods pose significant challenges for detection and prosecution.[54]

C Chronic Offenders

Chronic offenders are individuals who consistently engage in the creation of deepfakes in significant volumes. Their behaviour is influenced by deeply ingrained attitudes and strong integration into specific online subcultures that celebrate deepfake production. These offenders derive their identity and sense of status from their activities. Their motivations are multifaceted, often combining ideological commitment, psychological satisfaction, and a desire for social belonging.[55]

These chronic offenders have fully absorbed neutralisation techniques and the values of their subculture, which justify their actions. They tend to view criticism as a form of persecution and see legal repercussions as unjust oppression. This mindset makes them resistant to traditional deterrence strategies and unlikely to stop their behaviour voluntarily.[56]

From a sociological viewpoint, chronic offenders illustrate how differential association and immersion in a subculture can lead to the development of enduring criminal identities. Their extensive connections with like-minded individuals and prolonged exposure to rationalising narratives have significantly shaped their worldview. Effective intervention with chronic offenders must address not just individual behaviour but also disrupt the subcultural environments that sustain their criminal activities.

D Policy and Legal Implications

This typological framework has significant implications for criminal justice responses to deepfakes. A differentiated approach that recognises the distinct characteristics of occasional, professional, and chronic offenders would enable more effective prevention, intervention, and punishment strategies.[57]

---

[54] Lorenz Meinen et al, 'Deepfakes in Criminal Investigations: Interdisciplinary Research Directions for CMC Research' (2025) arXiv preprint.

[55] Sexualized Deepfake Abuse: Perpetrator and Victim Experiences (2025) Journal of Interpersonal Violence https://journals.sagepub.com/doi/full/10.1177/08862605251368834 accessed 05 October 2025.

[56] Sexually explicit deepfakes and the criminal law in NSW (NSW Parliament, 2025) https://www.parliament.nsw.gov.au/researchpapers/Documents/Sexually%20explicit%20deepfakes.pdf accessed 05 October 2025.

[57] F Romero-Moreno, 'Deepfake detection in generative AI: A legal framework for criminal liability' (2025) 13(2) Computer Law & Security Review.

ISSN: 2582-6433

For occasional offenders, general deterrence through awareness campaigns highlighting legal consequences, combined with accessible reporting mechanisms and swift removal of content, may prevent initial offences or encourage desistance. Educational interventions emphasising victim harm and ethical digital citizenship could reduce the pool of potential occasional offenders.

Professional offenders require targeted investigation and prosecution, utilising financial intelligence, international cooperation, and technological detection capabilities. Strategies should focus on disrupting the economic incentives and operational infrastructure that enable professional deepfake offending. Enhanced penalties for commercial deepfake production and distribution may increase the perceived risks in their cost-benefit calculations.

Chronic offenders necessitate comprehensive approaches that combine legal sanctions with platform intervention and community disruption. Dismantling or infiltrating online subcultures, removing influential members, and countering rationalising narratives can reduce the social reinforcement that sustains chronic offending. Long-term incarceration or significant restrictions may be appropriate for those who persistently create harmful deepfakes despite less severe interventions.

Across all offender types, technological solutions play a crucial role. Improved detection algorithms, mandatory watermarking of synthetic media, and authentication systems can reduce opportunity structures and increase guardianship. Platform accountability mechanisms, including liability for hosting deepfake content and obligations to implement preventive measures, address the capable guardianship deficit identified by Routine Activity Theory.

Legislative reforms should explicitly criminalise deepfake creation across various contexts, sexual, political, financial, and reputational, with penalties calibrated to harm severity and offender persistence. Jurisdictional cooperation frameworks are essential for addressing the transnational nature of deepfake offending. Civil remedies, including takedown rights, damages, and injunctive relief, should complement criminal sanctions to empower victims.

CONCLUSION AND SUGGESTIONS

This project critically examined the creation of malicious deepfakes from the perspectives of classical, psychological, and sociological criminological theories, along with an analysis of both Indian and international legal frameworks. The findings confirm that creating deepfakes is a deliberate act influenced by a complex interplay of individual rational calculations, psychological traits, social environments, and group identities, thereby supporting the proposed hypothesis. The rise of accessible deepfake technologies has expanded the pool of offenders,

ISSN: 2582-6433

while gaps in existing legal and regulatory frameworks, particularly in India, hinder effective prevention and prosecution.

To address these challenges, a multifaceted approach is necessary. Legal reforms should explicitly criminalise various forms of deepfake offences across sexual, political, and financial domains, with penalties which reflect different offender typologies, occasional, professional, and chronic. Enhanced technological measures, such as mandatory watermarking, improved detection algorithms, and increased platform accountability, must complement legislative efforts. Education and awareness campaigns aimed at occasional offenders can help reduce impulsive misuse, while strengthening law enforcement capabilities is crucial for targeted investigations into sophisticated deepfake crimes. Additionally, disrupting deviant online subcultures through community engagement and digital interventions is essential to curbing chronic offending.

Furthermore, fostering international legal cooperation is vital due to the transnational nature of deepfake offences. Investing in interdisciplinary research and developing proactive detection technologies will help keep pace with rapid advancements in synthetic media. Importantly, victim-centred remedies, including accessible reporting mechanisms and psychological support, should be incorporated into policy frameworks to mitigate harm. This integrative approach enhances both criminological theory and practical policy, emphasising that deepfakes represent not only technological threats but also evolving social phenomena that require coordinated legal, technological, and sociological responses.