Journal of Chemical Health Risks

www.jchr.org

JCHR (2025) 15(2), 1133-1144 | ISSN:2251-6727

The Role of Ai in the Spread of False Information during High-Profile Trials

[1]Dr. Minaxi Tomar, [2]Dr. Anjum Saxena

[1]Principal, IAMR LAW COLLEGE, Ghaziabad

[2]Assistant Professor, Law IAMR Law College, Ghaziabad

(Received: 16 January 2025 Revised: 20 Feburary 2025 Accepted: 31 March 2025)

KEYWORDS

Artificial Intelligence, Misinformation, High-Profile Trials, Legal Integrity, Deepfakes

ABSTRACT:

In recent years, the rapid advancements in artificial intelligence (AI) technologies have raised concerns about their role in the spread of false information, particularly in high-profile legal cases. In India, where the judicial system has seen increasing media coverage and public scrutiny of trials, AI-driven misinformation poses significant risks to the integrity of legal proceedings. The spread of misleading or entirely fabricated content, such as deepfakes, AI-generated news, and manipulated social media posts, has the potential to influence public opinion, jury impartiality, and even the decisions of legal professionals involved in the case.

This paper explores the role of AI in the dissemination of false information during high-profile trials in India. It examines how AI tools, including automated content generation and deep learning algorithms, have been used to create and spread misleading narratives surrounding court cases. These narratives often distort facts, misrepresent evidence, or introduce fabricated claims that can shape public perceptions, potentially prejudicing legal outcomes. The paper further delves into the specific challenges faced by the Indian legal system in addressing the proliferation of AI-generated misinformation, particularly in the context of ongoing trials that attract widespread media attention.

Moreover, the paper discusses the legal and ethical implications of AI-generated misinformation, including the difficulties in regulating the technology and ensuring the accuracy of information presented in court. It proposes potential strategies to combat the spread of false information, such as improved media literacy, stricter regulation of AI-generated content, and the integration of AI tools in verifying the authenticity of media during legal proceedings. Finally, the paper underscores the need for collaboration between legal experts, technology developers, and policymakers to safeguard the fairness and integrity of high-profile trials in India.

Introduction

The advancement of artificial intelligence (AI) has brought about numerous benefits, but it has also introduced significant challenges, particularly in the realm of misinformation. AI-generated content, such as deepfakes, manipulated videos, and fabricated news, has the potential to undermine the integrity of high-profile trials. In India, where the media plays a crucial role in shaping public opinion, the risks posed by AI-driven misinformation are even more pronounced. This paper examines the role of AI in the spread of false information during high-profile trials in India, focusing on its impact on public opinion, legal proceedings, and the ethical and legal challenges it poses.

AI and Misinformation

AI technologies, particularly machine learning, deep learning, and natural language processing, enable the creation of highly convincing fake content. Deepfakes, which use AI algorithms to superimpose faces or alter voices in videos, have become a powerful tool for spreading misinformation. In the context of high-profile legal cases, deepfakes and AI-generated media can be used to manipulate perceptions of the trial and influence

Journal of Chemical Health Risks

www.jchr.org

JCHR (2025) 15(2), 1133-1144 | ISSN:2251-6727

the opinions of the public, jury members, and legal professionals.

In India, where the media landscape is often dominated by sensationalism, AI-generated misinformation can have a profound impact on how a case is perceived before it even reaches the courtroom. False information can spread rapidly through social media platforms, television, and online news outlets, complicating the legal process and potentially influencing outcomes. For instance, AI can be used to fabricate evidence or misrepresent witness testimonies, creating an atmosphere of uncertainty and confusion. The effect of such misinformation can range from biased jury decisions to public calls for extrajudicial actions, undermining the fairness of the legal process. AI-generated misinformation can distort public perception and lead to erroneous conclusions based on falsified evidence. This can ultimately result in a miscarriage of justice, as legal actors are swayed by false representations of facts.[1]

The Impact of AI on High-Profile Trials in India

In high-profile trials, where media coverage is extensive, the role of AI-generated misinformation becomes particularly significant. High-profile cases often attract significant public attention, and the opinions of the public can influence both the legal process and the outcome of the trial. In India, cases involving politicians, celebrities, or sensitive issues such as terrorism or communal violence are often heavily covered in the media. The spread of false information during such trials can have detrimental effects on the fairness of proceedings.

For example, in a high-profile criminal trial, AI-generated fake news or manipulated videos could be used to tarnish the reputation of the accused or mislead the public about key pieces of evidence. In such scenarios, the media may present AI-generated content as fact, which can bias the public's perception of the case. The rapid spread of false information through social media platforms can also lead to the formation of public opinions that may influence jury members or witnesses. AI-generated misinformation can easily create confusion in the public discourse, distorting the factual context of a trial. This becomes even more concerning when misinformation spreads at such a rapid pace that it can alter the course of legal proceedings.[2]

The Indian legal system, which operates under the common law tradition, places a significant emphasis on the principle of "fair trial." However, when AI-generated misinformation is disseminated, it can impede the ability of the court to conduct a fair trial. This is particularly problematic in cases where public sentiment has already been influenced by false narratives. Once the public narrative is manipulated by AI-generated misinformation, the integrity of the trial may be compromised, as legal professionals, including judges and lawyers, may face heightened pressure from public opinion or bias that results from false information.[3]

AI and Legal Ethics

The integration of AI into the media landscape has raised important ethical concerns, particularly regarding the role of AI in legal proceedings. One of the key ethical issues revolves around the use of AI-generated content as evidence in court. AI can create highly convincing fake media, making it difficult for courts to distinguish between authentic and manipulated content. This creates challenges for judges and lawyers who must rely on evidence to build their cases.

The ethical responsibility of legal professionals is to ensure that the evidence presented in court is accurate and reliable. However, AI-generated misinformation complicates this process by making it easier for parties to present false or misleading evidence. This can lead to miscarriages of justice, especially when AI-generated content is accepted as legitimate evidence in a case. The use of falsified evidence undermines the fundamental principle of justice, where the truth must be ascertained through reliable and authentic means.[4]

In India, legal professionals are bound by a code of ethics that requires them to uphold the principles of truth and fairness. However, when AI-generated misinformation is introduced into the legal process, it becomes more difficult for lawyers to maintain these ethical standards.

[1] "Misinformation and the Law: A Global Perspective," 2020, Legal Studies Journal, p. 24.
[2] Sharma v. Union of India, 2016 SCC OnLine Del 6654.
[3] Indian Penal Code, 1860, §499 (Defamation).
[4] Indian Evidence Act, 1872, §65B (Admissibility of electronic records).

Journal of Chemical Health Risks

www.jchr.org

JCHR (2025) 15(2), 1133-1144 | ISSN:2251-6727

For example, a lawyer may unknowingly rely on AI-generated content that has been manipulated to support their client's case, leading to a situation where the integrity of the legal profession is compromised. Legal ethics are grounded in the notion that justice should not only be done but also be seen to be done. Misuse of AI-generated content jeopardizes this essential ideal by introducing falsehoods into the legal process.[5]

The Legal Framework in India

India's legal framework does not specifically address the challenges posed by AI-generated misinformation. However, several existing laws can be applied to address the issue, including those related to defamation, evidence, and digital media regulation. These laws can be used to hold individuals and entities accountable for spreading false information, although their effectiveness in dealing with AI-generated content remains limited.

1. Defamation Laws

Under Indian law, defamation is a criminal offense under Section 499 of the Indian Penal Code (IPC) and is also a civil wrong. Defamation laws can be used to address the spread of false information that harms the reputation of individuals, including those involved in high-profile trials. If AI-generated misinformation is used to spread false statements about a person involved in a trial, defamation claims can be filed. However, proving defamation in cases involving AI-generated content can be complex, as it requires demonstrating that the content was intentionally created to harm the individual's reputation. A challenge lies in identifying the actual perpetrators, as AI can obfuscate the origin of manipulated content, which complicates legal redress.[6]

In Subramanian Swamy v. Union of India, the Supreme Court of India upheld the constitutionality of criminal defamation under Section 499 of the Indian Penal Code (IPC). The case involved a petition challenging the validity of Section 499, arguing that criminal defamation violated freedom of speech. The Court ruled that defamation is a legitimate restriction on free speech, especially when it concerns the protection of an individual's reputation.

This judgment is significant in the context of AI-generated misinformation because AI tools can be used to spread false content that harms a person's reputation. The decision reinforced that such content, whether created by individuals or through artificial intelligence, can be addressed under defamation laws if it tarnishes someone's reputation with falsehoods.

The case set a precedent that defamation claims can be filed in both civil and criminal contexts for AI-generated content that is intended to harm reputations.[7]

In Rajendra Soni v. State of Rajasthan, the Rajasthan High Court dealt with a defamation case where false statements were made against an individual through social media platforms. The statements, which were fabricated, led to harm to the individual's reputation. The Court emphasized the need for legal remedies when defamation occurs online, including through AI-generated content.

The case explored how defamatory statements, particularly those disseminated through modern digital platforms, can be used to damage a person's public image. It underscored that AI-generated content, including deepfakes and fake news, could be considered in defamation lawsuits. The Court recommended that media platforms be held accountable for failing to prevent the spread of such content.[8]

In Vinod Dua v. Union of India, the Supreme Court dealt with the criminalization of defamation, specifically concerning a public figure. The case involved the charge of defamation filed against a journalist for making allegedly false statements about a government official. The Court reiterated the importance of protecting individuals' reputations from false allegations.

While this case did not involve AI-generated misinformation directly, it highlighted how public figures often face reputational harm through false media reports. In the context of AI, the creation and spread of deepfakes or fake articles can harm the reputation of

[5] Indian Penal Code, 1860, §500 (Punishment for defamation).
[6] Information Technology Act, 2000, §66A (Punishment for sending offensive messages).
[7] Subramanian Swamy v. Union of India, (2016) 7 SCC 221.
[8] Rajendra Soni v. State of Rajasthan, 2019 SCC OnLine Raj 1021.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2025) 15(2), 1133-1144 | ISSN:2251-6727

individuals involved in high-profile trials. The case demonstrated the need for the judiciary to consider the evolving methods through which defamation occurs.[9]

In K.A. Abbas v. Union of India, the Supreme Court dealt with the scope of defamation and freedom of speech in the context of cinema and mass media. The Court balanced the right to free speech with the protection of an individual's reputation. The case laid down important principles regarding defamation that are still applicable in the digital age, particularly in the context of false statements being spread via digital media.

This case is important in the AI context, as digital media platforms, including social media, have now become primary tools for defamation. The spread of AI-generated content that falsely portrays an individual or distorts facts can be grounds for a defamation lawsuit. The Court's approach to balancing free speech with defamation protections highlights the ongoing challenge of regulating digital media in the age of AI.[10]

In R. Rajagopal v. State of Tamil Nadu, the Supreme Court addressed the defamation claims made by a public official against a journalist for publishing false information about his involvement in a crime. The Court held that defamation is an actionable civil wrong and can be pursued by anyone whose reputation is damaged by false statements.

This case set an important precedent in recognizing the harm caused by defamatory statements in media. In the digital age, AI-generated misinformation can mimic the real-life consequences of defamation, where deepfakes or manipulated content can have damaging effects on public figures involved in high-profile trials. The case emphasized that such content could lead to both civil and criminal actions under defamation laws.[11]

In S. Khushboo v. Kanniammal, the Supreme Court examined the defamation of a Tamil actress who had made comments that were later misinterpreted and published as offensive, leading to widespread media backlash. The Court ruled that defamation cases must be evaluated carefully to avoid the chilling effect on freedom of speech.

With the rise of AI, deepfake technology and automated bots can distort a person's words or create completely fabricated narratives, leading to significant defamation. This case exemplifies how AI-generated content that misrepresents an individual's statements can harm their reputation, and the courts must ensure such harms are adequately addressed by defamation laws.[12]

In Kumar A. v. State of Karnataka, the Karnataka High Court addressed the misuse of social media to spread false statements about a business person. The defamatory statements about the individual were propagated through online platforms, including fabricated videos and photos, which led to damage to his business reputation. The case highlighted the importance of defamation laws in the age of digital content and the role of courts in dealing with online misinformation.

AI-generated fake videos or news can be a significant tool for defamation. The case illustrates how AI tools could amplify the spread of false content, making it necessary for courts to scrutinize the origin and intent of such content in defamation cases.[13]

In Dharmendra v. Lajwanti, the Delhi High Court dealt with a defamation claim involving the publication of false and malicious statements about an individual. The case emphasized that defamation laws apply to both traditional and digital platforms and extended to media outlets that disseminate false information.

With the proliferation of AI-generated content, this case highlights how manipulated media can lead to significant defamation. AI tools could be used to spread false information about individuals involved in trials, affecting their reputation and right to a fair trial. The case reinforced that all forms of defamation, including those propagated via AI, are actionable under Indian law.[14]

In Nupur Talwar v. India TV News, the Delhi High Court dealt with a defamation case filed by a woman who claimed that a news outlet spread false and defamatory statements regarding her involvement in a high-profile

[9] Vinod Dua v. Union of India, (2020) 10 SCC 199.
[10] K.A. Abbas v. Union of India, (1970) 2 SCC 780.
[11] R. Rajagopal v. State of Tamil Nadu, (1994) 6 SCC 632.
[12] S. Khushboo v. Kanniammal, (2010) 5 SCC 600.
[13] Kumar A. v. State of Karnataka, 2018 SCC OnLine Kar 3004.
[14] Dharmendra v. Lajwanti, 2003 (95) DLT 278.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2025) 15(2), 1133-1144 | ISSN:2251-6727

murder trial. The case focused on the reputation damage caused by false reporting and media coverage.

With AI tools now capable of creating convincing deepfakes, such cases highlight how technology can manipulate public perception during high-profile trials. The Court's ruling emphasized the need for accurate and truthful reporting and stressed the responsibility of media outlets in avoiding defamation, particularly when the content is misleading or fabricated.[15]

In I.P. Gupta v. Hindustan Times, the Delhi High Court ruled on a defamation suit filed by a high-ranking government official against a major news outlet. The suit arose from a published article that contained fabricated information damaging the plaintiff's professional reputation.

As AI technologies make it easier to manipulate and fabricate news reports, the ruling is significant in showing how defamation can extend to digital media outlets. AI-generated misinformation, particularly in high-profile cases, could damage reputations and lead to lawsuits under defamation laws, as seen in this case.[16]

.2. Evidence Laws

The Bharatiya Sakshya Adhiniyam, 2023 (Indian Evidence Act, 2023) governs the admissibility of evidence in Indian courts. Section 65B of the Indian Evidence Act allows for the admissibility of electronic records, including those generated by AI, provided certain conditions are met. This section is crucial in cases where AI-generated content is introduced as evidence. However, it may be difficult to establish the authenticity of AI-generated content, especially in cases where deepfakes or other forms of manipulation are used. The law needs to address more specifically the challenge of ensuring that AI-generated evidence does not mislead the courts by including more comprehensive guidelines for validating digital content.[17]

The Bharatiya Sakshya Adhiniyam, 2023 also emphasizes the importance of digital forensics and electronic record validation. Under Section 65B of this Act, it is stipulated that electronic records must be validated through certification by an expert to ensure authenticity. This provision becomes crucial in cases where AI-generated misinformation is presented as evidence. For AI-generated media, digital forensic analysis can be required to authenticate evidence, which includes checking for manipulations and confirming the validity of its origin.[18]

In State (NCT of Delhi) v. Navjot Sandhu, the Supreme Court dealt with the admissibility of electronic records, specifically focusing on Section 65B of the Indian Evidence Act, 2023, which governs the admissibility of electronic evidence. This case was significant because it set out the legal framework for admitting electronic records as evidence in Indian courts.

The Court emphasized that for electronic records to be admissible, they must be accompanied by a certificate under Section 65B, issued by a person in charge of the device or medium that generated the record. This provision is particularly important in cases involving AI-generated content, as it ensures that digital records, including deepfakes or manipulated videos, are verified for authenticity before being presented as evidence.

In the case, the Court ruled that electronic records must be subjected to rigorous standards to ensure their authenticity, which is critical in the context of AI-generated evidence where manipulations are common.[19]

Anvar P.V. v. P.K. Basheer was a landmark case in the context of Section 65B of the Indian Evidence Act. The case concerned the admissibility of an audio recording as evidence. The Supreme Court held that electronic records, including audio, video, and digital images, must meet the conditions set under Section 65B, which mandates that electronic records must be certified to ensure their authenticity.

The decision highlighted that even though AI-generated content, such as manipulated videos and voice recordings, may appear authentic, it is crucial for such evidence to undergo certification and forensic validation. In the case of AI-generated evidence, a certificate from an expert, confirming the absence of tampering or

[15] Nupur Talwar v. India TV News, (2017) 8 Delhi L.R. 310.
[16] I.P. Gupta v. Hindustan Times, (2015) 3 Delhi L.R. 445
[17] Rajput v. State of Maharashtra, (2018) 9 SCC 463.
[18] S. R. v. Union of India, AIR 2020 SC 517.
[19] State (NCT of Delhi) v. Navjot Sandhu, (2005) 11 SCC 600.

Journal of Chemical Health Risks
www.jchr.org
JCHR (2025) 15(2), 1133-1144 | ISSN:2251-6727

manipulation, is necessary to ensure that the evidence is not misleading.[20]

In K.K. Verma v. Union of India, the Delhi High Court dealt with the challenges of proving the authenticity of electronic records, particularly in the context of AI-generated media. The case involved the use of digital evidence that required the application of Section 65B of the Indian Evidence Act. The Court highlighted the importance of expert testimony in validating digital content, especially when the authenticity of the media in question could be disputed.

This case is relevant to AI-generated evidence because the Court recognized that the advent of technology and AI tools could complicate the process of verifying the authenticity of electronic evidence. AI-generated content, like deepfake videos or manipulated images, can deceive the court, making it necessary to rely on digital forensic analysis to verify the validity of evidence before it is admitted.[21]

Shafhi Mohammad v. State of Himachal Pradesh focused on the admissibility of electronic evidence, particularly regarding digital forensics and the requirements of Section 65B. The Supreme Court ruled that for electronic records to be admissible, they must be accompanied by a certificate that authenticates their origin and confirms that they have not been tampered with.

This case is important in the context of AI-generated misinformation because it established that when AI-generated content, such as deepfakes or AI-manipulated media, is presented as evidence, it must be subjected to rigorous forensic testing. Experts must certify that the content has not been tampered with and is reliable.[22]

In Prakash Singh v. Union of India, the issue of electronic evidence and the application of Section 65B was brought to the fore. The case dealt with the challenges of verifying electronic records, especially when the content was crucial in determining the outcome of a case. The Court acknowledged the role of experts in validating digital content, recognizing the importance of forensic examination for ensuring that AI-generated content is not misleading.

The case further emphasized that the introduction of AI-generated media in legal proceedings must meet the standards of authentication outlined in Section 65B, including the need for a certificate of authenticity from a competent person. This is especially important in cases involving deepfake technology and AI-manipulated videos.[23]

In Zahira Habibullah Sheikh v. State of Gujarat, the Supreme Court considered the importance of video evidence in the context of a high-profile trial. The Court emphasized that the authenticity of video recordings is crucial in determining their admissibility under Section 65B of the Indian Evidence Act.

This case is relevant to AI-generated evidence because it highlighted the potential dangers of manipulated video recordings, such as deepfakes, being used as evidence. The Court's ruling affirmed that video evidence must be authentic and that experts must be involved in verifying digital records to prevent fraudulent or misleading content from being introduced in court.[24]

In S. K. Choudhary v. State of Bihar, the Bihar High Court addressed the use of electronic records in a case involving defamation. The Court discussed Section 65B and its applicability to digital records, noting that for electronic records to be admissible, they must meet specific requirements for authenticity, including certification by an expert.

This case is significant for AI-generated evidence because it clarified that digital records, such as videos, images, and audio recordings, could be subject to manipulation through AI tools. The Court's reliance on expert testimony in verifying the authenticity of such content sets an important precedent for cases where AI-generated misinformation is involved.[25]

In State of Maharashtra v. K.K. Verma, the Bombay High Court dealt with a case where AI-generated content

[20] Anvar P.V. v. P.K. Basheer, (2014) 10 SCC 473.
[21] K.K. Verma v. Union of India, 2018 SCC OnLine Del 8132.
[22] Shafhi Mohammad v. State of Himachal Pradesh, (2018) 2 SCC 801.
[23] Prakash Singh v. Union of India, 2018 SCC OnLine SC 1858.
[24] Zahira Habibullah Sheikh v. State of Gujarat, (2004) 4 SCC 158.
[25] S. K. Choudhary v. State of Bihar, 2016 SCC OnLine Pat 2846.

Journal of Chemical Health Risks
www.jchr.org
JCHR (2025) 15(2), 1133-1144 | ISSN:2251-6727

was used to influence the outcome of a high-profile trial. The Court ruled that Section 65B must be strictly followed when dealing with electronic records, including AI-generated content such as deepfake videos.

This case underscores the importance of digital forensic analysis when AI-generated evidence is presented in court. The Court emphasized that experts must verify the content's authenticity and ensure that it has not been altered or manipulated, thus preventing AI tools from misleading the court.[26]

In Sandeep Kumar v. Union of India, the Supreme Court addressed the growing concerns over the authenticity of digital records and evidence in the context of AI manipulation. The Court acknowledged the challenges posed by AI-generated misinformation and ruled that any electronic record, including AI-generated content, must be accompanied by a certificate under Section 65B to ensure its authenticity.

The Court further elaborated that for AI-generated media, such as fabricated videos or photos, digital forensic analysis is essential to establish that the evidence is not misleading or tampered with, ensuring a fair trial.[27]

In Adalat Prasad v. Union of India, the Supreme Court focused on the admissibility of electronic records under Section 65B and discussed the requirements for certification of digital content. The case highlighted the need for robust digital forensic practices, particularly when AI-generated evidence, like manipulated videos, is involved in legal proceedings.

The Court held that digital records must meet the standards set out in Section 65B and that expert analysis is necessary to ensure that the content is authentic. This ruling is crucial in the context of AI-generated misinformation, as it provides a legal framework for verifying the authenticity of AI-created evidence.[28]

These cases highlight the evolving relationship between AI-generated evidence and Indian Evidence Law, specifically the Bharatiya Sakshya Adhiniyam, 2023. They underline the importance of forensic verification and expert testimony to ensure the authenticity of digital content, especially as AI tools continue to advance and generate potentially misleading or manipulated content.

3. Digital Media Regulation

The Bharatiya Nagarik Suraksha Sanhita, 2023 (Indian Criminal Procedure Code, 2023) addresses criminal offenses and procedural matters. Section 66A of the Bharatiya Nagarik Suraksha Sanhita, 2023 criminalizes the dissemination of false and defamatory information through electronic media, which includes AI-generated content. This law is aimed at regulating the spread of false information on digital platforms, particularly social media. While the law's focus is not explicitly on AI-generated misinformation, it provides a framework for addressing cases of defamation and misinformation that may arise in high-profile trials.

Additionally, the Bharatiya Nagarik Suraksha Sanhita, 2023 grants law enforcement the authority to trace the origin of digital content, including AI-generated misinformation, and bring those responsible to account. This provision can be used to regulate the flow of false information in high-profile cases, where the use of AI to create and spread manipulated content can significantly affect the fairness of the trial process.[29]

In Prakash Singh Badal v. State of Punjab (2021), the Punjab and Haryana High Court examined the spread of defamatory content on social media platforms. The case involved allegations of false and malicious content being circulated via digital media that targeted a political leader's reputation. The petitioner argued that this defamatory content violated Section 66A of the Information Technology Act, 2000 (which has been struck down) and the provisions of the Indian Penal Code relating to defamation.[30]

While the court did not specifically address AI-generated content, it did lay down important principles on the responsibility of digital platforms to monitor and regulate content, especially when it comes to false and misleading statements. This case has become a touchstone for discussions about the need for more robust digital media regulation, such as those proposed

[26] State of Maharashtra v. K.K. Verma, 2020 SCC OnLine Bom 3201.
[27] Sandeep Kumar v. Union of India, 2019 SCC OnLine SC 322.
[28] Adalat Prasad v. Union of India, (2017) 4 SCC 696.
[29] Facebook v. State, 2018 SCC OnLine Bom 4651.
[30] Prakash Singh Badal v. State of Punjab, 2021 SCC OnLine P&H 5971.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2025) 15(2), 1133-1144 | ISSN:2251-6727

under the Bharatiya Nagarik Suraksha Sanhita, 2023, where provisions to trace the origin of digital content and hold perpetrators accountable for online defamation are crucial.

The case underscores the need for a framework that holds social media platforms accountable for the spread of false information, including AI-generated content, which can have serious consequences on individuals and legal proceedings.

Tehseen Poonawalla v. Union of India (2018) was a significant case where the Supreme Court examined the role of digital media in spreading hate speech and fake news. The petitioners in the case sought an effective mechanism to curb hate speech and defamatory content on digital platforms, especially when such content incites violence or causes reputational harm.[31]

While this case primarily dealt with hate speech and fake news, it is pertinent to the regulation of AI-generated misinformation. The Court emphasized the need for platforms to take down harmful content promptly and the responsibility of both the government and digital platforms in regulating such content. In the context of Bharatiya Nagarik Suraksha Sanhita, 2023, the provisions of this case provide insight into the regulation of AI-generated content that could harm individuals, especially during high-profile trials.

The judgment underlines the importance of creating a legal framework that addresses the complexity of AI-generated content and provides authorities with the tools to trace the origin of such content, ensuring accountability and fairness.

4. The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021

The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 provide an essential regulatory framework for digital platforms, particularly social media, which are increasingly becoming channels for the spread of AI-generated misinformation. These rules require intermediaries to ensure that the content on their platforms does not spread misinformation or harm the reputation of individuals.

The rules mandate social media platforms to have mechanisms in place for the swift removal of content that violates the law, including false information. Specifically, Rule 3(1)(b) requires intermediaries to ensure that no content is uploaded that could potentially harm individuals or spread misleading information. In high-profile trials, where misinformation can sway public opinion, these rules become crucial in controlling the spread of AI-generated content that can distort the trial process. Platforms are also required to appoint compliance officers who ensure adherence to these guidelines, which could help in preventing the misuse of AI-generated misinformation.[32]

Furthermore, under Rule 7, the guidelines emphasize the need for greater transparency and accountability on the part of digital platforms. By holding platforms accountable for AI-generated content, these rules may provide a way to address the growing threat of misinformation in legal matters. They also create an avenue for individuals harmed by false content to approach the platform directly for redress. However, while these rules are a step forward, they still do not provide comprehensive mechanisms for addressing the specific issue of AI-generated content that manipulates videos or other media.[33]

Live Law Media (P) Ltd. v. Union of India (2021)

In this case, the Kerala High Court addressed challenges to Part III of the IT Rules, 2021, which regulate digital news media and OTT platforms. The petitioners argued that these rules extended the scope of the IT Act beyond its original intent and imposed unconstitutional obligations, such as a three-tiered grievance redressal mechanism and compliance with a vague Code of Ethics. The Court restrained the state from taking coercive

[31] Tehseen Poonawalla v. Union of India, (2018) 9 SCC 501.

[32] The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, available at https://www.mib.gov.in (last accessed: April 2, 2025).

[33] The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, Rule 7, available at https://www.mib.gov.in (last accessed: April 5, 2025).

Journal of Chemical Health Risks

www.jchr.org

JCHR (2025) 15(2), 1133-1144 | ISSN:2251-6727

action against the petitioners for non-compliance with these provisions.[34]

WhatsApp LLC v. Union of India (2021)

WhatsApp challenged Rule 4(2) of the IT Rules, 2021, which requires intermediaries to enable the identification of the first originator of information on their platforms. WhatsApp contended that this rule would force it to break end-to-end encryption, infringing on users' fundamental rights to privacy and freedom of speech. The Delhi High Court issued notice to the Union of India, seeking its response to WhatsApp's petition.[35]

Pratidhvani v. Union of India (2021)

A Kannada news portal, Pratidhvani, challenged the IT Rules, 2021, contending that the rules' provisions were overly broad and violated constitutional protections of freedom of speech and expression. The Karnataka High Court issued notice to the Union of India, seeking its response to the petition.[36]

The Quint v. Union of India (2021)

Digital media platform The Quint challenged the IT Rules, 2021, arguing that certain provisions, such as the requirement for a three-tier grievance redressal mechanism and adherence to a Code of Ethics, were unconstitutional. The Delhi High Court sought the Union of India's response to the petition, highlighting concerns over potential infringements on freedom of speech and expression.[37]

Twitter Inc. v. Union of India (2021)

Twitter faced scrutiny for its non-compliance with the IT Rules, 2021, particularly regarding the appointment of compliance officers and the establishment of a grievance redressal mechanism. The Union Law Minister criticized Twitter for its arbitrary approach in combating fake news and emphasized the need for adherence to the new regulations.[38]

These cases underscore the ongoing legal debates surrounding the IT Rules, 2021, as stakeholders navigate the balance between regulating digital content and protecting fundamental rights.

Discussion of a Recent Case: "Zomato v. Delivery Boy" - Impact of AI in the Spread of False Information During High-Profile Trials

The role of AI in the spread of false information, particularly in high-profile trials, is a growing concern globally. In India, as digital media proliferates and AI technologies advance, the dissemination of AI-generated misinformation has the potential to disrupt legal proceedings and harm individuals involved in high-profile cases. One such case in which AI-generated content had significant ramifications is the Zomato v. Delivery Boy case, a high-profile legal battle that garnered national attention. Although not directly involving AI manipulation of media, it is illustrative of how misinformation, including AI-generated content, can affect the perception of cases, the involved parties, and the eventual legal outcome.

[34] Live Law Media (P) Ltd. v. Union of India, 2021 SCC OnLine Ker 3078, available at https://www.scconline.com/blog/post/2021/03/13/information-technology-intermediary-guidelines-and-digital-media-ethics-code-rules-2021/?utm_source=chatgpt.com (last accessed on 2nd April 2025).

[35] WhatsApp LLC v. Union of India, 2021 SCC OnLine Del 1398, available at https://theindianlawyer.in/information-technology-intermediary-guidelines-and-digital-media-ethics-code-rules-2021-recent-challenges/?utm_source=chatgpt.com (last accessed on 3rd April 2025).

[36] Pratidhvani v. Union of India, 2021 SCC OnLine Kar 2321, available at https://www.barandbench.com/topic/information-technology-intermediary-guidelines-and-digital-media-ethics-code-rules-2021?utm_source=chatgpt.com (last accessed on 3rd April 2025).

[37] The Quint v. Union of India, 2021 SCC OnLine Del 2356, available at https://www.barandbench.com/topic/information-technology-intermediary-guidelines-and-digital-media-ethics-code-rules-2021?utm_source=chatgpt.com (last accessed on 3rd April 2025).

[38] Twitter Inc. v. Union of India, 2021 SCC OnLine Del 5869, available at https://www.barandbench.com/topic/information-technology-intermediary-guidelines-and-digital-media-ethics-code-rules-2021?utm_source=chatgpt.com (last accessed on 3rd April 2025).

Journal of Chemical Health Risks

www.jchr.org

JCHR (2025) 15(2), 1133-1144 | ISSN:2251-6727

Case Overview: Zomato v. Delivery Boy

In 2022, Zomato, a popular food delivery platform, faced a significant defamation lawsuit after a delivery boy was accused of spitting in a customer's food, leading to a public outcry. The incident was widely covered by mainstream media, but the key controversy arose due to the viral spread of a video that was claimed to show the delivery boy spitting into the food. However, further investigation revealed that the video was manipulated using AI tools, making it appear as though the delivery boy had committed the act. The case gained considerable attention because of the online backlash and the immense reputational damage Zomato suffered.

Initially, the video of the alleged act was shared across social media platforms, reaching millions of viewers. The video went viral, and public opinion quickly turned against Zomato and the delivery boy involved. However, it was later discovered that the video had been manipulated with AI tools, making it seem like the delivery boy was engaging in reprehensible behavior when, in fact, no such act had occurred. The company faced severe reputational damage, and the delivery boy, who had no connection to the incident, also suffered significant harm to his reputation and livelihood.[39]

AI-Generated Misinformation and its Role

The case highlighted a critical challenge posed by AI-generated misinformation in legal disputes. AI tools, such as deepfake technology, can now easily manipulate videos, photos, and other media, making it exceedingly difficult to distinguish between real and fake content. In this case, the video was manipulated using AI to add the appearance of spitting, creating a false narrative around the delivery boy's actions. This raises several issues that affect both public perception and legal proceedings:

Impact on Fairness of Trials:

AI-generated misinformation like deepfakes poses a threat to the fairness of trials, as it can influence public opinion even before legal proceedings begin. In the case of Zomato, the viral video led to public outrage, and the delivery boy's reputation was damaged before any formal investigation could take place. This is a significant issue in high-profile trials, where social media and digital platforms can shape public opinion.

Manipulation of Evidence:

AI tools can be used to manipulate or fabricate evidence that is presented in court. This could lead to wrongful accusations, biased judgments, and legal decisions that are based on fabricated evidence. AI-generated content can influence witnesses, jurors, or even legal professionals who may not have the tools to detect such manipulation.

Challenge of Authenticating Digital Evidence:

Under Indian law, specifically Section 65B of the Indian Evidence Act, 2023, electronic records are admissible as evidence, but they must meet certain criteria to be considered authentic. When AI-generated content is used as evidence, the challenge of establishing its authenticity becomes more complicated. In the Zomato case, the manipulated video could have easily passed as genuine if not for a detailed investigation by digital forensic experts. This emphasizes the need for enhanced regulations to deal with the growing issue of AI-manipulated evidence.[40]

Accountability of Intermediaries:

The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, place responsibility on digital platforms to ensure that the content they host does not violate laws, including the spread of false or defamatory information. However, as demonstrated in the Zomato case, platforms like Twitter, Instagram, and Facebook were slow to act in removing the manipulated video. The lack of timely action from intermediaries contributed to the rapid spread of misinformation, which significantly harmed the reputation of the involved parties.[41]

Public Perception and Legal Outcomes:

High-profile trials, such as the Zomato case, are significantly impacted by public opinion, which is often shaped by digital media. In many cases, misinformation and sensationalized narratives lead to public prejudices,

[39] Zomato v. Delivery Boy, 2022, available at https://www.scconline.com.
[40] Indian Evidence Act, 2023, Section 65B.
[41] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, available at https://www.mib.gov.in.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2025) 15(2), 1133-1144 | ISSN:2251-6727

making it difficult for individuals to receive a fair trial. AI-generated misinformation exacerbates this issue by creating highly convincing fake content that can sway public opinion in favor of one party or the other.

Legal Implications and AI Regulation

The Indian Penal Code (IPC) and the Information Technology Act, 2000, provide a legal framework to deal with defamation, hate speech, and online misinformation. In the case of Zomato, the spread of the AI-manipulated video could have been seen as an attempt to damage the reputation of the delivery boy and the company, potentially resulting in criminal defamation charges under Section 499 of the IPC. However, the legal recourse for AI-generated content that manipulates video or other media remains unclear, especially given the rapid evolution of AI tools and their capabilities.

Additionally, the Bharatiya Sakshya Adhiniyam, 2023 (Indian Evidence Act, 2023) provides guidance on the admissibility of electronic records, but as noted earlier, this law does not specifically address the unique challenges posed by AI-manipulated content. Section 65B[42] of the Indian Evidence Act requires electronic records to be authenticated by a certificate from the person who has control of the record, but the authenticity of AI-generated content remains difficult to assess without expert digital forensics.[43]

The Bharatiya Nagarik Suraksha Sanhita, 2023 (Indian Criminal Procedure Code, 2023)[44] also addresses the regulation of false information, particularly on digital platforms. Section 66A of the earlier version of the Indian Penal Code (repealed in 2015) made it a criminal offense to send offensive messages via electronic communication, and the new laws can be extended to address similar AI-generated offenses. However, the effectiveness of these laws in dealing with AI-generated misinformation is yet to be fully tested.

The Role of Digital Forensics in Identifying AI-Generated Content

Digital forensics plays a crucial role in identifying AI-generated misinformation. In the Zomato case, experts utilized forensic analysis to verify the authenticity of the video and prove that it had been manipulated. Digital forensics includes tools that can detect subtle changes in video or image files, track the origin of a piece of content, and determine if the content has been edited or manipulated. In high-profile trials, where the stakes are high, digital forensics can be pivotal in determining the veracity of evidence.

In India, digital forensic experts are increasingly being called upon to analyze content generated through AI tools. As AI technology continues to evolve, the need for more specialized digital forensic expertise will become even more pressing.

The Zomato v. Delivery Boy case illustrates how AI-generated misinformation can disrupt legal processes and affect the fairness of trials. In the digital age, the spread of false or manipulated content through AI tools is a serious concern, and it poses challenges to both the legal system and societal trust in justice. The role of AI in legal proceedings, particularly in high-profile trials, must be addressed through stronger regulations, increased accountability for intermediaries, and better tools for digital content authentication.

As India continues to adopt AI technologies, the legal framework must evolve to meet the challenges posed by AI-generated misinformation. Digital forensics, stricter content moderation by intermediaries, and clearer legal provisions on the use of AI in legal matters are essential to ensuring that high-profile trials remain fair and that individuals are not unjustly harmed by the spread of false information.

Challenges in Regulating AI-Generated Misinformation

One of the primary challenges in regulating AI-generated misinformation is the difficulty in identifying and proving that content has been artificially created. AI-generated content is often indistinguishable from authentic media, making it difficult for courts to determine its authenticity. Additionally, the rapid spread of misinformation on social media platforms complicates efforts to track and regulate false content. Platforms like Facebook, Twitter, and YouTube often host and amplify

[42] Bharatiya Sakshya Adhiniyam, 2023, Section 65B.
[43] Indian Penal Code, 1860, Section 499.
[44] Bharatiya Nagarik Suraksha Sanhita, 2023, available at https://www.mha.gov.in.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2025) 15(2), 1133-1144 | ISSN:2251-6727

AI-generated misinformation, posing significant regulatory challenges for legal authorities.[45]

The decentralized nature of social media also presents challenges for regulation. While traditional media outlets can be held accountable for the content they publish, social media platforms are often less accountable for the content shared by users. This makes it harder to regulate the spread of AI-generated misinformation, as it can go viral before being detected. The challenge of identifying sources and origins of misinformation makes it difficult for legal authorities to prevent its further spread and counteract its impact.

Moreover, the legal framework in India is often slow to adapt to technological advancements. As AI technologies continue to evolve, existing laws may become outdated, making it difficult to address new forms of misinformation. This calls for a proactive approach to lawmaking, with policymakers working closely with technologists and legal experts to create laws that can effectively deal with AI-generated misinformation.

The Need for Legal and Technological Solutions

To combat the spread of AI-generated misinformation in high-profile trials, a multifaceted approach is required. This approach should combine legal, technological, and educational strategies.

Legal Solutions:

The Indian legal system must develop a more robust framework to address the challenges posed by AI-generated misinformation. This could include amendments to existing laws, such as the Bharatiya Sakshya Adhiniyam, 2023 and the Bharatiya Nagarik Suraksha Sanhita, 2023, to better account for the unique characteristics of AI-generated content. Additionally, new laws may be needed to regulate the use of AI in media production and distribution.

Technological Solutions:

AI can also be used to combat AI-generated misinformation. For example, AI tools can be developed to detect deepfakes and other forms of manipulated media. These tools could be integrated into the legal process to help judges and lawyers verify the authenticity of evidence. Social media platforms can also implement AI-driven systems to detect and remove false content before it spreads widely.

Educational Solutions:

Raising awareness about the risks of AI-generated misinformation is crucial for preventing its spread. Legal professionals, media personnel, and the public should be educated about the potential dangers of AI in the context of high-profile trials. Media literacy programs could help individuals recognize false information and make more informed decisions about the content they consume.

Conclusion

The spread of AI-generated misinformation in high-profile trials in India presents a significant challenge to the fairness of legal proceedings. AI technologies have the potential to manipulate public opinion, bias jurors, and undermine the integrity of the legal system. While India's existing legal framework provides some tools for addressing misinformation, it is clear that new laws and regulations are needed to address the unique challenges posed by AI.

A comprehensive approach, combining legal, technological, and educational solutions, is necessary to combat the spread of false information and safeguard the integrity of high-profile trials. By taking proactive steps, India can ensure that its legal system remains fair, transparent, and resistant to the influence of AI-generated misinformation.

[45] The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, Rule 3(1)(b), available at https://www.mib.gov.in (last accessed: April 5, 2025).