Peer - Reviewed & Refereed Journal

The Law Journal strives to provide a platform for discussion of International as well as National Developments in the Field of Law.

DISCLAIMER

EDITORIAL
TEAM

Raju Narayana Swamy (IAS) Indian Administrative Service officer

Dr. Raju Narayana Swamy popularly known as Kerala's Anti-Corruption Crusader is the All India Topper of the 1991 batch of the IAS and is currently posted as Principal Secretary to the Government of Kerala. He has earned many accolades as he hit against the political-bureaucrat corruption nexus in India. Dr Swamy holds a B.Tech in Computer Science and Engineering from the IIT Madras and a Ph. D. in Cyber Law from Gujarat National Law University. He also has an LLM (Pro) (with specialization in IPR) as well as three PG Diplomas from the National Law University, Delhi- one in Urban Environmental Management and Law, another in Environmental Law and Policy and a third one in Tourism and Environmental Law. He also holds a post-graduate diploma in IPR from the National Law School, Bengaluru and a professional diploma in Public Procurement from the World Bank.

Dr. R. K. Upadhyay

Dr. R. K. Upadhyay is Registrar, University of Kota (Raj.), Dr Upadhyay obtained LLB , LLM degrees from Banaras Hindu University & Phd from university of Kota.He has succesfully completed UGC sponsored M.R.P for the work in the ares of the various prisoners reforms in the state of the Rajasthan.

<u>Senior Editor</u>

Dr. Neha Mishra

Dr. Neha Mishra is Associate Professor & Associate Dean (Scholarships) in Jindal Global Law School, OP Jindal Global University. She was awarded both her PhD degree and Associate Professor & Associate Dean M.A.; LL.B. (University of Delhi); LL.M.; Ph.D. (NLSIU, Bangalore) LLM from National Law School of India University, Bengaluru; she did her LL.B. from Faculty of Law, Delhi University as well as M.A. and B.A. from Hindu College and DCAC from DU respectively. Neha has been a Visiting Fellow, School of Social Work, Michigan State University, 2016 and invited speaker Panelist at Global Conference, Whitney R. Harris World Law Institute, Washington University in St.Louis, 2015.

<u>Ms. Sumiti Ahuja</u>

Ms. Sumiti Ahuja, Assistant Professor, Faculty of Law, University of Delhi,

Ms. Sumiti Ahuja completed her LL.M. from the Indian Law Institute with specialization in Criminal Law and Corporate Law, and has over nine years of teaching experience. She has done her LL.B. from the Faculty of Law, University of Delhi. She is currently pursuing Ph.D. in the area of Forensics and Law. Prior to joining the teaching profession, she has worked as Research Assistant for projects funded by different agencies of Govt. of India. She has developed various audio-video teaching modules under UGC e-PG Pathshala programme in the area of Criminology, under the aegis of an MHRD Project. Her areas of interest are Criminal Law, Law of Evidence, Interpretation of Statutes, and Clinical Legal Education.

<u>Dr. Navtika Singh Nautiyal</u>

Dr. Navtika Singh Nautiyal presently working as an Assistant Professor in School of law, Forensic Justice and Policy studies at National Forensic Sciences University, Gandhinagar, Gujarat. She has 9 years of Teaching and Research Experience. She has completed her Philosophy of Doctorate in 'Intercountry adoption laws from Uttranchal University, Dehradun' and LLM from Indian Law Institute, New Delhi.

Dr. Rinu Saraswat

Associate Professor at School of Law, Apex University, Jaipur, M.A, LL.M, Ph.D,

Dr. Rinu have 5 yrs of teaching experience in renowned institutions like Jagannath University and Apex University. Participated in more than 20 national and international seminars and conferences and 5 workshops and training programmes.

Dr. Nitesh Saraswat

E.MBA, LL.M, Ph.D, PGDSAPM
Currently working as Assistant Professor at Law Centre II, Faculty of Law, University of Delhi. Dr. Nitesh have 14 years of Teaching, Administrative and research experience in Renowned Institutions like Amity University, Tata Institute of Social Sciences, Jai Narain Vyas University Jodhpur, Jagannath University and Nirma University.
More than 25 Publications in renowned National and International Journals and has authored a Text book on Cr.P.C and Juvenile Delinquency law.

Subhrajit Chanda

BBA. LL.B. (Hons.) (Amity University, Rajasthan); LL. M. (UPES, Dehradun) (Nottingham Trent University, UK); Ph.D. Candidate (G.D. Goenka University)

Subhrajit did his LL.M. in Sports Law, from Nottingham Trent University of United Kingdoms, with international scholarship provided by university; he has also completed another LL.M. in Energy Law from University of Petroleum and Energy Studies, India. He did his B.B.A.LL.B. (Hons.) focussing on International Trade Law.

ABOUT US

WHITE BLACK LEGAL is an open access, peer-reviewed and refereed journal providededicated to express views on topical legal issues, thereby generating a cross current of ideas on emerging matters. This platform shall also ignite the initiative and desire of young law students to contribute in the field of law. The erudite response of legal luminaries shall be solicited to enable readers to explore challenges that lie before law makers, lawyers and the society at large, in the event of the ever changing social, economic and technological scenario.

With this thought, we hereby present to you
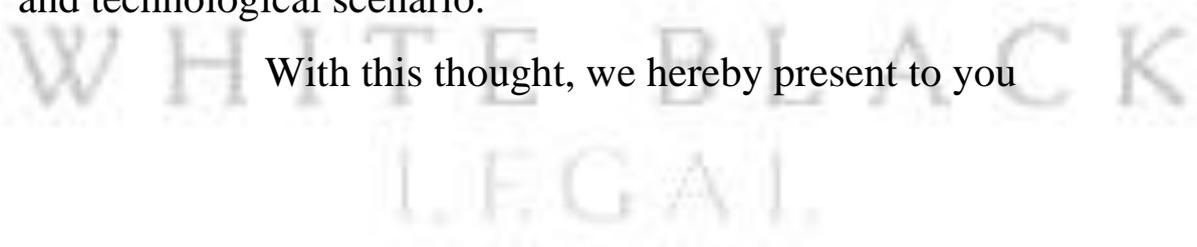
DEEPFAKES AND THE LAW: NAVIGATING THE THIN LINE BETWEEN SATIRE AND DEFAMATION

AUTHORED BY - E. KIRTHIKA

5th Year, BA. LLB (Hons)

Saveetha School of Law

Imagine waking up one morning to find a video of yourself circulating online giving a controversial speech you never delivered, wearing clothes you don't own, and standing in a place you've never been. The voice, the gestures, the expressions all uncannily yours. Friends are calling, news outlets are picking it up, and your reputation is unravelling before you've had a chance to deny it.

This isn't a scene from a sci-fi thriller it's the unsettling reality made possible by **deepfake technology**, a form of synthetic media created using artificial intelligence to generate realistic yet entirely fabricated videos, images, or audio. What began as a tool for digital creativity and entertainment has quickly evolved into a double-edged sword one that threatens the very foundation of truth, trust, and accountability in the digital world.

With applications ranging from political misinformation and identity theft to satire and parody, deepfakes occupy a legally ambiguous and ethically charged space. The challenge lies in determining when this technology crosses the line from protected free expression to harmful defamation and what legal safeguards are necessary to prevent its abuse while preserving its legitimate use.

The rise of generative artificial intelligence (AI) has introduced a powerful but controversial tool **deepfakes**. These hyper-realistic, fabricated videos, images, and audio clips can convincingly mimic a person's likeness and voice, blurring the line between reality and manipulation. While deepfakes are increasingly used in entertainment, satire, and creative media, their darker applications such as defamation, identity theft, political misinformation, and the erosion of public trust pose serious legal and ethical challenges.

In India, as in many parts of the world, the legal system is still catching up to the rapid technological advancements in AI. Although existing laws related to defamation, privacy, and intellectual property offer partial protection, the absence of AI-specific legislation leaves significant grey areas in enforcement, accountability, and redressal.

Satire or Defamation? Understanding the Legal Debate

A central concern in the deepfake discourse is distinguishing between satire, which enjoys constitutional protection as a form of free speech, and defamation, which can lead to legal liability. The key lies in the intent behind the content and the impact it has on the subject.

- **Satire** is a legitimate form of expression that uses humour, irony, or exaggeration to critique public figures, institutions, or societal norms. It often serves as a powerful tool in democratic discourse.

- **Defamation**, however, involves false representations that harm an individual's reputation or cause public misperception. A deepfake that misrepresents a person in a misleading, derogatory, or offensive manner may fall within the realm of defamation, especially if it causes reputational damage or public outrage.

Satire or Slander? Where Humour Meets Harm

Now picture another scenario. A popular comedy show features a skit where an actor mimics a well-known politician fumbling through a press conference, mispronouncing words, and nervously dodging questions about corruption. The audience laughs, social media erupts with memes, and the politician issues a statement calling it "defamatory." But was it really?

This is where satire enters the conversation. Satire is a recognized form of artistic and social expression that uses humour, irony, or exaggeration to critique public figures, institutions, or societal behaviour. It doesn't aim to deceive, but to provoke thought, expose flaws, and spark conversation often by holding up a mirror to power.

Unlike deepfakes meant to mislead or impersonate, satire relies on context and intent. It may imitate or caricature real people, but its purpose is clear: to entertain or criticize, not to impersonate or cause reputational damage under false pretences. Yet in today's digital world where parody, misinformation, and AI-driven manipulation often overlap the boundary between satire and defamation is growing increasingly difficult to define.

Defamation in Action: When Deepfakes Go Too Far

Consider this example: A deepfake video surfaces online showing a respected doctor accepting a bribe and making unethical remarks about patient care. The video goes viral, leading to public outrage, suspension from work, and irreparable damage to the doctor's reputation. Only days later, it's revealed that the video was fake AI-generated using publicly available images and voice samples. In this case, the deepfake is not satire it is **defamation**, as it falsely depicts criminal behaviour and severely harms the individual's professional and personal life.

Such examples highlight why legal clarity and enforcement are urgently needed. The stakes are no longer hypothetical they're real, reputational, and potentially irreversible.

Women, Consent, and Non-Consensual Deepfake Pornography

One of the darkest and most urgent frontiers in deepfake misuse is non-consensual pornographic content. Numerous cases have emerged globally where AI-generated videos depict women often public figures or even minors in explicit situations they were never part of. These deepfakes not only humiliate but also traumatize victims, who are left battling social stigma, legal hurdles, and digital permanence.

In India, such content falls under Section 67A of the IT Act[1], which criminalizes the transmission of sexually explicit material. However, victims often struggle with platform inaction, delayed removal, and procedural red tape. Unlike satire or political misinformation, these deepfakes have no conceivable defence they are designed to exploit, intimidate, and silence. Calls for urgent reforms include fast-track takedown protocols, psychological support for victims, and stronger punitive provisions for creators and distributors of such content. For women and marginalized groups, the battle against deepfakes is not just about reputation it's about reclaiming agency and safety in digital spaces.

Case Spotlight: Anil Kapoor vs Simply Life & Ors

A recent and significant development in the Indian legal landscape is the case of *Anil Kapoor vs Simply Life & Ors[2]*. The veteran actor approached the Delhi High Court seeking protection

---

[1] Ransom, Shashank Shekhar Ashish. "Ethical &Amp; Legal Implications of Deep Fake Technology: A Global Overview." *Proceeding International Conference on Science and Engineering* 11, no. 1 (February 18, 2023): 2226–35. https://doi.org/10.52783/cienceng.v11i1.398.

[2] Khurana, Tarun, and Tarun Khurana. "Understanding Personality Theft: Anil Kapoor Vs Simply Life." Khurana and Khurana, February 19, 2025. https://www.khuranaandkhurana.com/2025/02/19/understanding-the-relevance-of-the-anil-kapoor-vs-simply-life-india-ors-case/.

against the unauthorized and harmful use of his identity including his name, image, voice, signature phrase "jhakaas," and digitally manipulated deepfakes that paired his likeness with actresses like Katrina Kaif and the late Sridevi.

The court took strong cognizance of how the misuse extended to merchandise, fake autographs, AI-generated cartoons, and even domain names that misled the public. Kapoor's counsel argued that such unauthorized exploitation violated his right to publicity, reputation, and livelihood, citing international precedents like *Bette Midler v. Ford Motor Co[3]*. and *Vanna White v. Samsung Electronics[4]*. The court agreed, stating that fame should not strip individuals of their dignity or privacy and that celebrities are entitled to protect their identities from unfair commercial or reputational harm. It prohibited the use of Kapoor's persona across various mediums, including GIFs, stickers, t-shirts, and domain names. This case clearly illustrates how deepfakes, under the guise of humour or satire, can cause reputational damage and financial loss, especially to public figures. The judgment is a landmark in recognizing the legal boundaries of satire and the rights of individuals over their personality in the AI era.

Consent, Control, and the Right to Be Forgotten

Beyond reputation, deepfakes raise important questions about consent and autonomy. Can someone's past videos, images, or interviews be repurposed indefinitely without their approval? In an era where data permanence has overtaken privacy, the "right to be forgotten" is gaining traction globally.

EU's General Data Protection Regulation (GDPR) already acknowledges this right, allowing individuals to request the erasure of personal data. India's forthcoming Digital India Act is also expected to include similar provisions. When applied to deepfakes, this could mean that individuals may one day claim the legal right to have unauthorized digital manipulations of themselves erased whether defamatory, comedic, or otherwise. Until then, the absence of individual consent remains a critical vulnerability in the digital age.

---

[3] Studicata. "Midler V. Ford Motor Co. - Case Brief Summary for Law School Success," July 2, 2024. https://studicata.com/case-briefs/case/midler-v-ford-motor-co/.
[4] Kozinski, J. "THE OVERPROTECTION OF CELEBRITY: A COMMENT ON WHITE V. SAMSUNG ELECTRONICS AMERICA, INC." *HOFSTRA LAW REVIEW*, 1994, 729–71. https://www.hofstralawreview.org/wp-content/uploads/2014/05/27_22HofstraLRev7291993-1994.pdf.

Political Deepfakes: A New Threat to Democracy

The political sphere is increasingly falling prey to the disruptive power of deepfakes, raising alarming concerns about their potential to distort democratic processes and mislead the public. With elections held across numerous countries in 2024, malicious actors have turned to AI-generated audio and visuals to sway opinions, suppress voter turnout, and fabricate endorsements.

One striking instance occurred ahead of a key U.S. primary election, where numerous voters in New Hampshire received a pre-recorded call mimicking the voice of President Joe Biden[5]. The message, falsely urging citizens to refrain from voting in the primary and instead wait until the general election, was later revealed to be an AI-generated deepfake. Such manipulation of political messaging, designed to create confusion and voter apathy, marks a dangerous precedent.

In another case, former U.S. President Donald Trump circulated[6] an AI-generated image falsely suggesting that popular singer Taylor Swift had endorsed him a claim entirely untrue. Though the image was exaggerated and cartoonish, its viral spread highlighted how easily such content can blur the lines between parody and misinformation, especially when shared without context.

These examples reflect a worrying trend: the weaponization of deepfakes to misrepresent political figures, mislead the electorate, and erode trust in democratic institutions. As deepfake technology grows more accessible and convincing, the urgent need for legal safeguards, media literacy, and platform accountability becomes all the more critical.

---

[5] AP News. "Fake Biden Robocall Being Investigated in New Hampshire | AP News," January 23, 2024. https://apnews.com/article/new-hampshire-primary-biden-ai-deepfake-robocallf3469ceb6dd613079092287994663db5.

[6] TOI World Desk. "US Elections: Did Donald Trump Use AI Images to Falsely Claim Taylor Swift's Endorsement?" *The Times of India*, August 19, 2024. https://timesofindia.indiatimes.com/world/us/us-elections-did-donald-trump-use-ai-generated-images-to-falsely-claim-taylor-swifts-endorsement/articleshow/112621159.cms.

India's Regulatory Response: Cracking Down on AI-Powered Misinformation

Recognizing the growing threat of deepfakes, the Ministry of Electronics and Information Technology (MeitY)[7] has issued a formal advisory to digital intermediaries, directing them to strengthen compliance with the IT Rules. The focus lies particularly on Rule 3(1)(b), which mandates platforms to actively prevent the hosting or circulation of harmful and deceptive content, including AI-generated misinformation.

The advisory requires platforms to clearly convey prohibited content policies to users not only through terms of service, but also at critical interaction points like registration, login, and content uploads. Users must be regularly reminded of the consequences of sharing manipulated or impersonated content, including deepfakes, and be made aware of penal provisions under the Indian Penal Code and IT Act.

Importantly, intermediaries are now obligated to act swiftly in detecting and removing misleading AI-generated content. The government has notified that failure to comply could result in legal action, and further regulatory tightening is on the horizon. This reflects a growing consensus that unchecked deepfakes are not only a violation of personal rights but a direct threat to digital safety and public trust.

Judicial Scrutiny and Government Response to Deepfake Threats

The growing misuse of deepfakes has now reached the doors of the judiciary, with the Delhi High Court hearing petitions that demand urgent regulation of this technology. A significant plea, filed by journalist Rajat Sharma, emphasized how deepfakes when left unchecked can distort public discourse, mislead voters, and irreparably damage reputations[8]. In its response, the Ministry of Electronics and Information Technology (MeitY) submitted a status report detailing the rise in AI-driven scams, political misinformation, and impersonation, especially during elections. While the report noted that India's existing legal framework the IT Act, IT Rules, and Bharatiya Nyaya Sanhita can address harmful deepfakes, it also acknowledged the

---

[7] "MeitY Issues Advisory to All Intermediaries to Comply With Existing IT Rules.," n.d. https://www.pib.gov.in/PressReleaseIframePage.aspx?PRID=1990542.

[8] Kohli, Tushar, Tushar Kohli, and Verdictum. "Verdictum." Verdictum, February 28, 2025. https://www.verdictum.in/court-updates/high-courts/injunction-likeness-personality-infringement-rajat-sharma-v tamara-doc-1147-of-2024-1561798.

need for better enforcement, technical tools, and user education to effectively counter such content.

The report further stressed the absence of a clear legal definition of "deepfake," which complicates regulatory efforts and enforcement. It highlighted that malicious actors often bypass detection tools like watermarking and metadata tracking, while victims of deepfake abuse are yet to be formally consulted. In response, the Delhi High Court directed[9] the formation of a committee to explore regulatory solutions and advised the government to incorporate victim testimonies and stakeholder inputs. The consensus among experts is clear: instead of stifling creative expression, regulations must target the malicious use of AI, ensuring that satire remains protected while defamation and digital deception are held legally accountable.

Media Literacy & Digital Hygiene: Empowering Citizens in the Age of Deepfake

"While legal and technical solutions are evolving, media literacy among digital citizens remains the first line of defence against the manipulation of truth in the age of AI."

While governments, courts, and tech platforms work to regulate deepfake misuse, an equally crucial pillar in this fight is public awareness. Legal safeguards and AI detection tools are essential, but they can only go so far without an informed and vigilant user base. Media literacy understanding how digital content is created, shared, and potentially manipulated is the first line of defence in the age of synthetic media. Citizens must be equipped to critically assess the authenticity of videos and audio, verify sources before sharing, and report harmful content to the appropriate platforms or authorities.

Digital hygiene practices such as avoiding unverified downloads, fact-checking viral claims, and being cautious of emotionally provocative content can drastically reduce the reach and impact of malicious deepfakes. Government campaigns, civil society efforts, and school curriculums must integrate digital literacy as a foundational skill. As deepfakes become more accessible and convincing, empowering individuals to detect and respond to them responsibly is not just advisable it's essential.

---

[9] PTI, and PTI. "Delhi High Court Directs Centre to Nominate Panel Members to Examine Deepfake Menace." Deccan Herald, November 23, 2024. https://www.deccanherald.com/india/delhi-high-court-directs-centre-to-nominate-panel-members-to-examine-deepfake-menace-3289124.

Beyond the Blur: Deepfakes as Tools for Empowerment and Innovation

Amid growing concerns around deception and defamation, it's easy to overlook that deepfake technology when divorced from malice can serve as a catalyst for progress. At its core, synthetic media is not inherently unethical; its value depends on intent and context. While regulation is essential to prevent misuse, dismissing deepfakes entirely risks stalling innovation in several transformative domains.

In digital heritage preservation, for instance, deepfake technology has been used to recreate the voices and faces of historical figures, allowing museums and educational institutions to offer immersive storytelling experiences that go beyond static exhibits. These reconstructions, guided by archival accuracy, breathe life into history for new generations without compromising factual integrity.

In mental health and therapy, some researchers are experimenting with personalized AI avatars that simulate empathetic conversations or help trauma survivors confront difficult memories in controlled virtual settings. These applications, still in early stages, suggest that the emotional realism of deepfakes can aid therapeutic healing when used responsibly.

Deepfakes also open doors for those historically excluded from mainstream media and content creation. Independent filmmakers with limited budgets are leveraging AI to generate background characters, de-age actors, or recreate lost footage once the preserve of high-end studios. For users with speech impairments or neurodivergent communication styles, deepfake voice synthesis paired with adaptive AI can offer new tools for expression, bridging gaps that conventional interfaces ignore.

Moreover, in humanitarian contexts, voice cloning has been used to deliver emergency messages in local dialects during crises, improving reach and trust among communities. Similarly, AI-driven simulations can train first responders using hyper-realistic scenarios, enhancing preparedness without placing anyone at real risk.

In a world increasingly reliant on digital interaction, deepfake technology if carefully governed can serve not as a threat, but as a means of amplifying voices, preserving culture, and personalizing communication in ways we've only begun to explore. The challenge lies not in condemning the tool, but in ensuring that creativity and accountability evolve in tandem.

India's Multi-Pronged Strategy Against Deepfakes: Beyond Legal Frameworks

To ensure a safe and accountable digital environment, the Government of India has taken a multi-layered approach to counter deepfake threats. While the Information Technology Act, 2000 and the IT Rules, 2021 form the legal foundation, enforcement efforts have expanded through both advisory mechanisms and technical preparedness.

The Ministry of Electronics and Information Technology (MeitY) has conducted several consultations with industry stakeholders, urging digital intermediaries to comply with their due diligence obligations particularly regarding misinformation and synthetic media. Platforms are expected to proactively remove deepfake content, ensure content moderation policies are clearly communicated to users, and inform them of legal consequences under the IT Act and IPC.

In addition to legal instruments, MeitY has issued multiple advisories aimed at tackling malicious AI-generated content. These efforts are complemented by initiatives such as the **Grievance Appellate Committees (GAC)**, which offer users the ability to challenge platform decisions on unlawful content via www.gac.gov.in.

On the cybersecurity front, the **Indian Computer Emergency Response Team (CERT-In)** has taken an active role. It regularly publishes alerts and advisories about threats arising from AI-based tools, including deepfakes, phishing, and social engineering. Its **November 2024 advisory** outlined safety measures to counter deepfake risks, while its **Cyber Swachhta Kendra** continues to provide free malware detection and removal tools.

Additionally, CERT-In promotes public awareness through annual initiatives like **Cyber Security Awareness Month**, **Safer Internet Day**, and **Cyber Jagrookta Diwas**, educating both citizens and institutions on safe digital practices.

The **Ministry of Home Affairs (MHA)** also plays a crucial role through the **Indian Cyber Crime Coordination Centre (I4C)** and the **National Cyber Crime Reporting Portal** (cybercrime.gov.in)[10], where victims can report cybercrimes, including deepfakes and AI-

---

[10] "Government of India Taking Measures to Tackle Deepfakes," n.d. https://www.pib.gov.in/PressReleasePage.aspx?PRID=2119050.

driven fraud.

Together, these mechanisms reflect a growing recognition that legal frameworks alone are not enough **technical vigilance, citizen awareness, and platform accountability** must go hand-in-hand to combat the evolving menace of deepfakes in India.

The Future of AI Regulation: From Liability to Transparency

As deepfakes continue to evolve, the regulatory focus must shift from reactive takedowns to proactive transparency. Future frameworks may require AI developers and content creators to embed "disclosure layers" or watermarks into synthetic content. These identifiers could help platforms and users quickly distinguish manipulated media from authentic ones. Additionally, there is growing debate around "algorithmic liability" whether developers of generative AI models should bear legal responsibility for harms caused by their tools. This becomes crucial when bad actors misuse open-source AI frameworks to create harmful content. A system of graded responsibilities from creators to intermediaries may soon form the backbone of AI regulation. In this evolving legal environment, maintaining user trust, safeguarding fundamental rights, and preserving technological progress will require not just rules, but adaptive governance models that evolve with innovation itself.

Conclusion

The debate over deepfakes is not merely about legality it's about redefining our relationship with reality itself. At a time when "seeing is believing" no longer holds true, truth itself becomes vulnerable to manipulation, misrepresentation, or even extinction. Deepfakes challenge the very foundations of law, ethics, consent, and accountability in the digital age. India's approach a blend of legal, technical, and citizen-focused responses offers a promising start, but cannot remain static. As deepfake technology grows more convincing and accessible, so too must our vigilance, education, and enforcement. We need robust laws that distinguish malice from creativity, tech tools that ensure traceability, platforms that act swiftly, and citizens who are empowered to question what they consume. In the end, the future of deepfakes will not be defined by the technology itself, but by the values we choose to uphold in confronting it. If governed wisely, deepfakes can fuel innovation, expression, and accessibility. If left unchecked, they risk becoming the most potent form of digital deceit our world has ever seen.