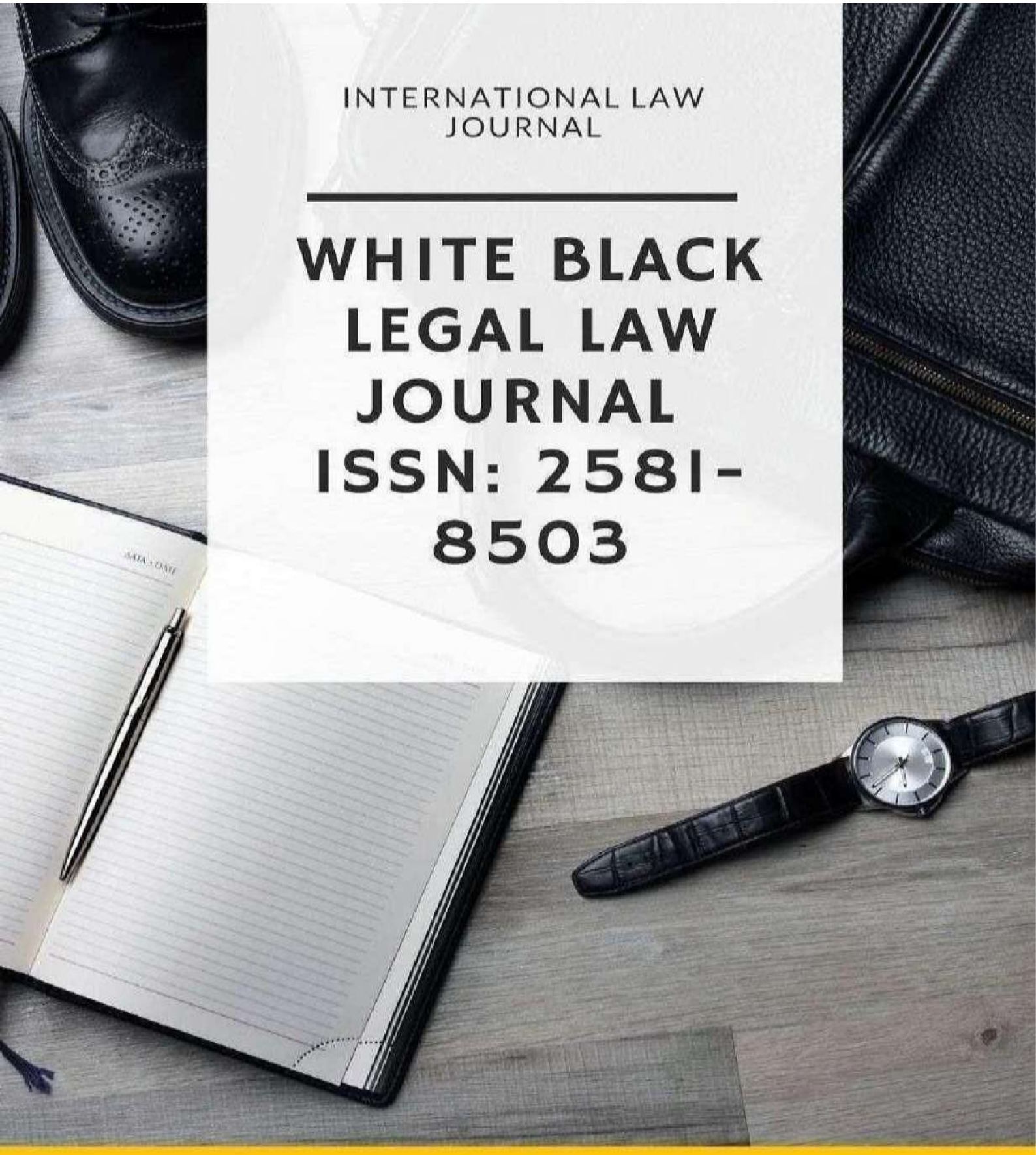

INTERNATIONAL LAW JOURNAL

WHITE BLACK LEGAL LAW JOURNAL

ISSN: 2581-8503

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 diploma in Public a professional 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.

Senior Editor

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.

Ms. Sumiti Ahuja

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.

Dr. Navtika Singh Nautiyal

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: EMERGING CRIMINALITY DISCOURSE AGAINST WOMEN

AUTHORED BY - SNEHAL UPADHYAY

ABSTRACT

In the era of technological advancements and introduction of Artificial Intelligence (AI) has led to emergence of various cyber-criminal activities the most prominent amongst all is creation of Deepfakes. The term Deepfakes connotes creation of an image or recording that has been convincingly altered or manipulated to misrepresent someone as doing or saying something that was not said or done by that particular person. In a study conducted by an AI firm Deeptrace[1] in the year 2019 it was found that most of the deepfakes were pornographic and 99% of the content involved women. It's an emerging threat for the society and is a new form of discourse as most of the time it's used for committing crime against women by making fake videos, obscene photographs and pornographic contents which causes image based abuse. In the present scenario India lacks on having an independent legislation on regulation of AI. Although, Deepfakes might fall under the purview of Indian Penal Code, 1860 and Information and technology Act, 2000 but still there is a legal vacuum, deepfakes are so complicated that it needs to have an Independent legislation. Through this research paper, the researcher wants to dwell into identifying the pitfalls with regard to Deepfakes impact on increasing crime against women. Further, the paper will also analyse the legal status of deepfakes in India with a comparative study of that to USA. Lastly, the research paper would offer suggestions on the loopholes and the gap which the existing laws has and would propagate on why there is a need of independent legislation for dealing with the issues on misuse of AI.

KEYWORDS: Deepfakes, Artificial Intelligence, pornography, cyber-criminal activities, Women, Image based abuse.

[1] Most Deepfakes Are Porn, and They're Multiplying Fast Most Deepfakes Are Porn, and They're Multiplying Fast | WIRED

RESEARCH OBJECTIVE AND METHEDOLOGY

RESEARCH OBJECTIVES:

The objective of this research paper is to:

1. To understand how deepfake technology is a threat and is an emerging criminality
2. To point out the new forms of crime committed against women and pitfalls with the help of AI.
3. To analyse the legal vacuum in the existing laws and emphasizing on the need to have regulation to deal with crime committed by AI separately.
4. To draw an analysis with USA's legal framework on crimes committed against women through AI
5. To suggest how the laws should be made in order to combat this type of criminality.

RESEARCH METHEDOLOGY

The researchers conducted a narrative review to explore the adverse impacts experienced by women due to technology and deepfakes, avoiding the oversimplification of results into standard metrics. They synthesized a range of sources, including emerging scholarly literature, news articles, and peer-reviewed studies from technological sectors. Searches were performed on Google Scholar and further other databases, as well as on the Google search engine, using a combination of search terms related to cyber threats, gendered image-based and video-based content, media abuse, harassment, and deepfakes. These terms were chosen to capture the various ways women are harmed online, including through specific types of deepfake content and platforms. By combining these search terms with Boolean operators, the researchers aimed to gather a comprehensive understanding of the threats posed by digital technologies and deepfakes to women.

LITRATURE REVIEW

This literature review delves into the harm inflicted upon women by technology, particularly focusing on the impact of deepfakes. It commences with a concise overview of cyber abuse targeting women, followed by an examination of literature concerning image-based sexual abuse against women. Lastly, the review delves into deepfakes and their associated harm.

1. INTRODUCTION TO DEEPFAKES

Defining deepfakes

With the onset of technological advancement around the globe making our lives easier but somewhere becoming a threat for our society Artificial Intelligence (AI) has come into existence. Deepfakes which is a type of an AI has recently become a major issue to combat and to prevent the targeted groups. Deepfake technology can be understood as it harnesses the power of artificial intelligence, specifically through the use of advanced algorithms like generative adversarial networks (GANs), to craft convincing synthetic media, including images, videos, and audio recordings. The primary aim is to fabricate media that closely resembles real content but with alterations. Deepfake technology relies on two key methodologies: deep learning and GANs. Deep learning, a subset of machine learning, employs algorithms modeled after the human brain's neural networks to analyze vast datasets, finding applications in fields like computer vision, natural language processing, and robotics. GANs, on the other hand, constitute a specific architecture within deep learning, utilizing two neural networks – a generator and a discriminator – to generate new data resembling a given dataset. The generator fabricates synthetic samples, while the discriminator evaluates their authenticity compared to real samples from the dataset. Through an adversarial training process, the generator strives to produce increasingly realistic data that can deceive the discriminator, iteratively improving until highly convincing synthetic media is achieved. Deepfake technology is such kind of technology where tone, modulation and facial expression can be adjusted in a single frame and distinguished features of 2 or more individual can also be combined for the requirement having high quality.[2]

Creation of Deepfakes

Deepfakes are generated through a machine learning approach called generative adversarial networks (GANs). GANs consist of two neural networks: a generator and a discriminator. These networks are trained on a comprehensive dataset containing real images, videos, or audio recordings. The generator produces synthetic data, such as images, resembling the real data in the training set. The discriminator evaluates the authenticity of this synthetic data and offers feedback

[2] Ian J. Goodfellow et al., "Generative Adversarial Nets", 27TH ANNUAL CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS (NIPS 2014), MONTREAL, 203 (Dec. 8, 2014), https://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf

to the generator for refinement. This iterative process continues until the generator generates highly realistic synthetic data that's challenging to differentiate from real data. These deepfakes can be utilized in various ways for video and image manipulation, including face swapping, altering attributes like hairstyle or color, transferring facial expressions, and creating entirely synthetic material based on real-world appearances.

- ***Offences caused using Deepfakes***

With the emerging use of deepfakes crime against women has been increased with the misuse of this very technology. With the misuse of deepfake technology by cybercriminals it has raised significant concern, especially regarding its exploitation in crimes, particularly those targeting women online. It has brought a new criminality discourse and which is increasing day by day. Further, it has been normalized so much that the deepfakes photographs and pictures gets keep on circulating on social media. Further, many illegal websites including pornographic sites too use deepfake technology to create pornographic content, revenge porn, non-consensual porn and many other different forms of harassment and deepfake is one more in the list like issue of threats or sale of fake obscene material are the emerging new form of criminality against women.

2. MISUSE OF DEEPFAKES

- ***Image based sexual abuse***

The emergence of deepfake technology has introduced a new dimension to the discourse surrounding cyber abuse, particularly concerning women. Traditionally, discussions have centered on image-based sexual abuse (IBSA), where intimate content is distributed without consent, predominantly affecting women and perpetrated by men. The ease of generating and disseminating harmful media through various online platforms has heightened concerns about the normalization and popularization of IBSA. The inception of many websites like IsAnyoneUp.com in 2010 exemplifies this trend, encouraging the submission of non-consensual content and personal information of women. The advent of advanced digital technologies, particularly AI tools and deepfake applications, has further exacerbated the threat of non-consensual sharing and exploitation. Perpetrators can now engage in abusive behavior anonymously across platforms, evading accountability for their actions. The psychological toll on victims of IBSA is profound, with reports of helplessness, distress, and disruption to social well-being. Women subjected to

IBSA often experience trauma, anxiety, and feelings of vulnerability, exacerbating concerns about safety, especially in cases linked to stalking and violence.[3]

Cyber sextortion, a subset of IBSA, poses unique challenges, leveraging explicit content to coerce victims into compliance with various demands. Unlike traditional forms of abuse, cyber sextortion occurs exclusively online, with perpetrators wielding threatening content to manipulate victims. Victims, fearing the potential dissemination of intimate material, may endure emotional distress and resort to drastic measures to protect their privacy, such as withdrawing from online spaces, changing contact information, and even relocating.

- ***Deepfakes and Women***

The emergence of deepfake technology presents a concerning trend in the realm of criminal discourse against women. With the proliferation of multimedia content in cyberspace, facilitated by technological advancements, the realism of digital impersonation has reached new heights. Women, historically vulnerable to various forms of exploitation but with the technological advancement now they counter vulnerability and exploitation in media and online spaces as well these digital manipulations, often non-consensual and sexually explicit in nature, serve as tools to compromise women's identities, damage their reputations, and enforce silence.

The distribution of deepfakes perpetuates an online environment where women's images are commodified for the pleasure of predominantly male users. Despite the prevalence of nonconsensual deepfake content, legal recourse for victims remains limited. Existing laws and regulations pertaining to deepfakes are often insufficient in providing adequate protection or avenues for redress. Compounding this issue is the anonymity of deepfake creators, who exploit various online platforms to disseminate manipulated content, making it challenging to identify and hold them accountable.

The detrimental impact of deepfake attacks on women extends beyond the digital realm, affecting their social, professional, and personal lives. Victims often experience significant psychological distress upon discovering their media content has been deepfaked, reporting feelings of

[3] JENNIFER LAFFIER AALYIA REHMAN, Deepfakes and Harm to Women, Journal of Digital Life and Learning, P. 5

humiliation, violation, fear, and powerlessness. The effects of deepfake attacks can include online and offline harassment, damaged professional reputations, and personal violation, leading to severe mental health issues and even suicidal ideation.

Furthermore, deepfake videos not only perpetuate harmful stereotypes and objectification of women but also normalize the notion that women can be abused and violated at will within online spaces. The stigmatization attached to women's bodies and actions exacerbates the collateral consequences and reputational damage inflicted upon deepfake victims. In essence, deepfake technology perpetuates a cycle of victimization and exploitation, underscoring the urgent need for comprehensive legal frameworks and technological safeguards to combat this emerging form of criminality against women.

3. UNDERSTANDING THE USE OF DEEPFAKES IN COMMITING CRIMES AGAINST WOMEN

Deepfakes have emerged as a significant concern within the realm of criminality, particularly in their exploitation of women. Initially, the genesis of deepfake technology found its foothold in pornography, with sensity.ai reporting a staggering 96% of deepfakes being of a pornographic nature, accumulating over 135 million views on adult websites[4]. This prevalence underscores the disproportionate targeting of women and increasing crime against women although men can also fall victim to such manipulative practices. However, it's crucial to acknowledge that women are often primary targets due to the enduring societal objectification that reduces them to mere sexual commodities. The ramifications extend beyond the digital realm, manifesting in emotional anguish, financial repercussions, and even loss of employment. Such exploitation not only violates individual privacy but also perpetuates harmful stereotypes and diminishes women's autonomy.

Revenge porn constitutes another facet of deepfake exploitation, as exemplified by the harrowing experiences shared by Zimbabwean women on BBC's The She Word[5]. The repercussions of being subjected to image-based sexual assault are severe, leading to societal ostracization, educational

[4] Ashish Jaiman, "The danger of deepfakes", THE HINDU, 1st Jan 2023, https://www.thehindu.com/sci-tech/technology/the-danger-of-deepfakes/article66327991.ece.

[5] "The She World", BBC WORLD SCIENCE TV, 19th Dec 2019, https://www.bbc.co.uk/programmes/p07xs7qs.

hindrances, and unemployment. Moreover, the psychological toll inflicted on victims, including anxiety, PTSD, and substance abuse, exacerbates their suffering. Intriguingly, studies reveal disparities in societal responses to male and female victims, with men often receiving more favorable treatment from law enforcement. This discrepancy underscores pervasive societal attitudes that may inadvertently perpetuate victim-blaming narratives, reinforcing the notion that women could have somehow prevented their victimization.

While other forms of image-based sexual abuse typically involve consensual creation, deepfakes differ in that they are generated without the consent of the individuals depicted. With the advancement of technology, creating deepfake images of celebrities, acquaintances, or anyone with accessible imagery has become increasingly feasible. Many perpetrators of deepfake creation seek personal sexual gratification without intending to distribute the imagery. This shift in demographic is evident, with a majority of respondents in a 2019 poll expressing a desire to create deepfakes of individuals they know personally, such actions are ethically permissible since these deepfakes are merely virtual images generated from publicly available information, akin to a sexual fantasy. However, this perspective overlooks the ethical implications of creating deepfakes, which constitute a violation of the sexual privacy of the victims and infringe upon their autonomy. Deepfakes represent a form of virtual sexual coercion and abuse, allowing individuals to digitally undress and exploit women without their consent. This violation not only contradicts the expectation that sexual activity should be based on mutual consent but also undermines individuals' ability to control the disclosure of their sexuality or gender, crucial aspects of identity formation. The repercussions of such violations can be profound, leading to lasting effects on victims' ability to trust others and develop intimate relationships.

Celebrity porn further underscores the pervasive nature of deepfake exploitation, as famous personalities like Scarlett Johansson and Gal Gadot find themselves unwilling participants in digitally manipulated sexual scenarios. However, the threat extends beyond the realm of celebrity, as anyone could potentially fall victim to having their likeness superimposed onto pornographic material. Recently, Deepfaked pictures of actress Rashmiska Mandana were also circulated in the social media and there have been many more celebrities who have become a target of deepfakes including morphed photographs and videos. This highlights the indiscriminate nature of deepfake

exploitation and the urgent need for comprehensive safeguards to protect individuals from such violations.[6]

Sextortion, also known as "eWhoring," is a form of extortion that involves threatening to distribute sexually explicit photos, whether obtained voluntarily or not. This threat typically involves blackmail, where the victim is coerced into fulfilling certain conditions under the threat of the explicit material being made public. This can include harming the victim physically, targeting their loved ones, damaging their property, or tarnishing their reputation.[7] While sextortion doesn't involve the sophisticated technology of deepfake for creating manipulated images, it can still be highly effective due to the sensitive nature of the material involved.

It's crucial for individuals to be aware of deepfake technology and its potential dangers. Deepfake technology can be used not only to create morphed and fake explicit pictures but also to manipulate audio, including mimicking voices. This poses risks in various sectors, including voice recognition systems used by government organizations for identification verification and by banks for financial transactions.

- ***The threat- Deepfake Pornography***

Deepfakes, propelled by advancements in AI technology such as General Adversarial Networks (GANs), have revolutionized the landscape of image and video manipulation, enabling the creation of hyper-realistic synthetic content indistinguishable from reality. Unlike traditional forms of image manipulation, deepfakes benefit from vast data sets and sophisticated algorithms, facilitating their widespread dissemination and commodification across various online platforms. This unprecedented speed, scale, and capacity with which deepfakes are generated and circulated present multifaceted challenges that demand a nuanced understanding of the harms they inflict.

At the macro-level, deepfake pornography emerges as a significant threat predominantly targeting women, perpetuating their sexual objectification and degradation within society. Indian women,

[6] K Melville, “The insidious rise of deepfake porn videos — and one woman who won't be silenced”, ABC NEWS, 30 Aug. 2019, 11:00AM, https://www.abc.net.au/news/2019-08 30/deepfake-revenge-porn-noelle-martin-story-of-image-based-abuse/11437774 (accessed 1 Mar. 2024).
[7] Konrad, K. A., & Skaperdas, S. (1998), ECONOMICA, 65, pp 461-477

particularly those in the entertainment industry, are disproportionately affected, with a majority of deepfake pornographic content featuring them. Moreover, politicians and journalists often become victims of sexual harassment and abuse through non-consensual deepfake pornography, highlighting the weaponization of this technology to silence dissent and discredit individuals. On a micro-level, victims of deepfake exploitation endure profound emotional, psychological, and financial repercussions. Beyond facing intimidation and bullying, victims are subjected to sextortion and revenge porn, leading to reputational damage that jeopardizes their career prospects and personal well-being. The ramifications extend beyond the individual, impacting their families and professional networks, underscoring the pervasive threat posed by deepfakes to societal norms and gender equality[8].

Addressing the multifaceted harms of deepfakes requires a holistic approach that transcends traditional legal avenues. Mitigation strategies must shift from solely relying on state institutions to engaging non-governmental organizations and platform economies. By decentralizing risk mitigation efforts, the accessibility and affordability of remedies can be enhanced while mitigating the risk of political misuse. Moreover, such an approach acknowledges the diverse interests at play in regulating deepfake pornography, paving the way for collaborative solutions that prioritize the protection of individuals' dignity and safety in the digital age.

4. REFERENCE TO CASES THROUGHOUT THE WORLD

As the advent of deepfakes has brought different types of criminalities which has been discussed in earlier chapters like non- consensual pornography and sextortion. These egregious violations of privacy and dignity have garnered significant media attention, highlighting the urgent need for effective legal measures to address this growing threat.

The case studies of individuals like Rana Ayyub, an investigative journalist, and Noelle Martin, a high school student, highlight the real-world consequences of deepfake victimization. Ayyub faced a barrage of obscene deepfake videos and threats after making political comments, leading her to practice self-censorship and withdraw from digital spaces. Martin's experience with deepfake

[8] Roger Brownsword, Eloise Scotford, and Karen Yeung, 'Law, Regulation, and Technology: The Field, Frame, and Focal Questions', in Roger Brownsword, Eloise Scotford, and Karen Yeung (eds), The Oxford Handbook of Law, Regulation and Technology, (Oxford Publications, 2017), p. 3-38.

attacks resulted in death threats, rape threats, and lasting social and psychological trauma, affecting her education and comfort in public settings.

Further examples of abound of individuals, predominantly women, falling victim to the insidious manipulation of deepfake technology. In India, cases of revenge pornography have left countless women exposed and vulnerable. From the horrifying betrayal of trust by uploading private photographs to the internet to the dissemination of intimate images by estranged partners, these incidents underscore the ease with which deepfake technology can be weaponized to inflict harm. The rapid advancement of deepfake software, exemplified by the emergence of apps capable of generating nude photographs within minutes, further exacerbates the threat posed by such malicious activities.

Even the recent social media trending video of two girls playing Holi together inside the metro faced backlash from the viewers for being obscene and hence, made Noida Police to find both the girls and imposed fine on both of them. Later, the Delhi Metro (DMRC) revealed that the video was a deepfake and hence, the girls were not playing Holi inside the metro[9]. Deepfakes have become so real that its very hard to identify from naked eye that what is real and what is reel.

One particularly distressing case involved a man in Ahmadabad who maliciously portrayed his wife and sister-in-law as prostitutes on social media platforms, leveraging private chats and naked images to perpetrate his defamation[10]. Similarly, an ex-husband in the same location stooped to despicable depths by emailing personal images of himself and his wife, taken during their honeymoon, to his sister-in-law and her spouse. These reprehensible acts not only violate the privacy and dignity of the victims but also underscore the urgent need for robust legal frameworks to combat such heinous offenses[11].

The prevalence of revenge porn in Indian society underscores the gravity of the issue and the

[9] Reel is real'? Delhi Metro (DMRC) claims clip deepfake, THE TIMES OF INDIA (28th March, 2024) http://timesofindia.indiatimes.com/articleshow/108831003.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst (accessed on 15.04.2024)

[10] TNN, "Separated Husband Posts Intimate Pictures", THE TIMES OF INDIA, (July 13th 2017) https://timesofindia.indiatimes.com/city/ahmedabad/separated-husband-posts-intimate pictures/articleshow/59568022.cms (accessed on April 12th 2024)

urgent need for intervention. However, the emergence of deepfake technology threatens to exacerbate this already pervasive problem, as the creation and dissemination of fake pornographic material become increasingly facile. Indeed, the very fabric of societal morals and decency is under siege from the proliferation of deepfakes, necessitating swift and decisive action on the part of lawmakers and law enforcement agencies. In light of these challenges, it is imperative that comprehensive laws and procedures be enacted to address the scourge of deepfake-related crimes against women. Legal measures must not only criminalize the creation and dissemination of non-consensual pornography but also provide avenues for redress and rehabilitation for victims because there is a threat to the mental health of the victims as well as the gravity of the offence is similar to that of any offence done physically. Moreover, concerted efforts are needed to raise awareness about the dangers of deepfake technology and empower individuals to protect themselves against such malicious activities. Only through a multi-faceted approach encompassing legislative, technological, and societal interventions can we hope to mitigate the impact of deepfakes on women and safeguard their rights and dignity in the digital age.

5. LEGAL STATUS OF TACKLING CRIMES RELATING TO DEEPFAKES IN INDIA

The emergence of deepfake technology has exacerbated the landscape of criminality against women, as evidenced by the increasing number of cases reported to the National Crime Records Bureau (NCRB). According to the NCRB Report of 2020[11], a staggering total of 2,756 cases were registered for offenses against women, encompassing cyber pornography, cyber stalking, cyberbullying, defamation, and indecent representation of women. Alarmingly, a significant portion of these cases, totaling 1,463, were attributed to cyber pornography, with revenge being cited as a primary motive. At the very recent stage NCRB has not brought the recent data on deepfakes crimes against women.

Deepfakes, with their capacity to manipulate video and images with unprecedented realism, have become a potent tool for perpetrating crimes against individuals, particularly women. These improper applications of AI-powered technology include identity theft, blackmail through manipulated video or images (commonly referred to as sextortion), and internet theft, among

[11] NCRB (National Crime Records Bureau) Report 2020

others. Such offenses not only violate the privacy and dignity of victims but also pose significant threats to their safety and well-being.

- ***Recognition of deepfakes as a crime under Indian Laws***

The emergence of deepfakes and AI-related crimes in India has identified significant gaps in our legal system as we fail to have a proper legislation to deal with crimes related to use of AI but there have been instances where people have got remedies under existing legislation making deepfakes to fall under their ambit.

The first and the foremost law which is hampered by the use of deepfakes is the Right to Privacy which is guaranteed under the Article 21 of the constitution. Creating deepfakes hampers one's privacy as in deepfake there's use of pictures of two different individuals to create a new image first to harm the reputation of the targeted individual. Therefore, it hampers Article 21 of the constitution.

The Information Technology Act, 2000 (IT Act), Section 66E addresses deepfake crimes involving the unauthorized capture, publication, or transmission of a person's images in mass media, thereby violating their privacy. Offenders may face imprisonment of up to three years or a fine of ₹2 lakh. Similarly, Section 66D punishes individuals who maliciously use communication devices or computer resources for impersonation or cheating, with penalties of up to three years imprisonment and/or a fine of ₹1 lakh.

Furthermore, Sections 67, 67A, and 67B of the IT Act empower authorities to prosecute individuals for publishing or transmitting deepfakes containing obscene or sexually explicit content. The IT Rules complement these provisions by prohibiting the hosting of content that impersonates another person and requiring social media platforms to promptly remove artificially morphed images upon notification. Failure to comply with these requirements may result in the loss of 'safe harbor' protection for social media companies.

The Indian Penal Code, 1860 (IPC), offers additional avenues for addressing cybercrimes associated with deepfakes. Sections 509, 499, and 153(a) and (b) enable prosecution for offenses

such as insulting the modesty of a woman, criminal defamation, and spreading hate on communal lines. In the case of *Sunilakhya v. HM Jadwet*[12] it was held that in criminal law, the act of intending to harm a person's reputation is a crucial aspect of a defamation offense. Deepfakes, encompassing visual depictions, fall within the realm of defamation. Notably, the Delhi Police Special Cell has invoked Sections 465 and 469 in the case of Rashmika Mandanna's viral deepfakes[13], registering an FIR against unknown persons for forgery and forgery to harm the reputation of a party. The notion of obscenity typically pertains to material that contravenes community standards and public decency, eliciting repulsion and lascivious interest. Under the Indian legal framework, Section 292 of the Indian Penal Code (IPC) explicitly addresses the sale, distribution, importation, or exportation of such obscene material, prescribing punishment for offenders. Image-based sexual abuse ('IBSA') and 'revenge pornography' are regulated as non-consensual intimate image distribution ('NCIID') under ss. 354A, 354C, 354D, 509 of IPC and ss. 66E, 66C, 67/67A of the IT Act.

Moreover, the Copyright Act of 1957 can be leveraged to address deepfakes involving copyrighted material. Section 51 prohibits the unauthorized use of another person's property, including images or videos, on which the latter enjoys exclusive rights.

Even though our existing laws can be used to make people liable for making deepfakes, but it's still not enough. As commission of crime reads the gravity of intention of doing a particular crime and the gravity of harm that particular person goes to should be matched. Therefore, the existing legislations are not a direct remedy but an indirect remedy until we do not have a proper legislation. As with time and advancement of technology new forms of criminality will emerge, nobody would have ever thought that a modesty of a women could be harmed by the use of deepfakes but it is happening, the offenders mind is clear enough to harm the reputation and modesty of women by circulating and selling obscene deepfaked pictures.

- ***Countering persisting Legal Vacuum***

The emergence of deepfakes has exposed significant deficiencies in existing legal frameworks,

[12] Sunilakhya v. HM Jadwet, AIR 1968 Cal 266.

[13] "Regulating deepfakes and generative AI in India" THE HINDU (December 04, 2023) Regulating deepfakes and generative AI in India | Explained - The Hindu (accessed on 14/4/2024)

prompting calls for comprehensive regulatory reform. Shehnaz Ahmed, a fintech lead at the Vidhi Centre for Legal Policy in Delhi[14], emphasizes the inadequacy of current laws, which were not crafted with emerging technologies like deepfakes in mind. While acknowledging the urgency to address recent high-profile cases, Ahmed cautions against piecemeal legislative amendments, urging instead for a broader examination of India's regulatory approach to emerging technologies such as AI.

Ahmed advocates for a regulatory framework grounded in a thorough market study to assess the diverse harms inflicted by AI technology. She stresses the importance of a robust enforcement mechanism, underscoring that effective implementation of laws requires institutional capacity. Highlighting a gap in the existing IT Rules, Ahmed notes their focus on addressing harm after illegal content has been uploaded, rather than prioritizing preventive measures such as raising awareness among users about the presence of morphed images.

With the provided reference it can be observed that the issue remains prevalent even after having laws that means the gravity of offences which are being committed by the use of AI has much more gravity for which the punishment and the remedies provided by the existing laws are not worthy and not satisfactory. Virtual world is very dangerous and the worst part is it is accessible by all young, adult and old everyone can access it from anywhere at any time hence, it exploits the mindset specially that of teenagers, the use to deepfakes and the curiosity of the teenage groups give them exposure and might make their mind to work on a wrong direction by learning to create deepfakes and then committing crimes. Therefore, there is no protection or precaution that the existing laws offer, which again brings an open question to the law makers that something which is blooming and might have a possibility of causing possible harm then why there is no legislation and hence calls for an urgent need of a legislation which completely deals with the laws and rules related to crimes and threats that AI can cause because AI is free and is accessible by all.

As it is freely accessible to all, pornographic content of any women can be created freely and circulated to pornographic sites. Even though the victim might not have shared her pictures but a

[14] "Regulating deepfakes and generative AI in India" THE HINDU (December 04, 2023) Regulating deepfakes and generative AI in India | Explained - The Hindu (accessed on 14/4/2024)

nude picture can be easily deepfaked and circulated in mass. There have been many victims and as stated the report of NCRB last updated in the year 2020 which is like 4 years back from now then it's evident that the figures might have increased massively.

These observations underscore the imperative for a holistic approach to legislative reform, one that not only addresses the immediate challenges posed by deepfakes but also anticipates and mitigates potential future threats arising from advancements in AI technology. By incorporating preventive measures, enhancing enforcement mechanisms, and ensuring greater accountability, India can forge a regulatory framework better equipped to safeguard against the pernicious impacts of deepfakes on women and society at large.

- ***Need of a detailed regulation on AI related crimes***

The proposed Digital India Act, 2023 draft[15], discussed by the Ministry of Electronics and Information Technology, reflects a growing recognition of the need to combat revenge porn and cyberbullying through appropriate legal measures. The inclusion of provisions for quality testing of AI-based technologies underscores the importance of regulating digital content to mitigate risks associated with deepfake technology.

A recent Suo Moto action[16] by the Supreme Court of India further exemplifies the judiciary's proactive stance in combating digital crimes, particularly those involving child pornography and sexual violence. The Ministry of Home Affairs' affidavit highlights the urgency of eradicating such content from online platforms, underscoring the legislative responsibility to draft and adopt guidelines to curb these heinous offenses. The court's order granting law enforcement bodies the authority to collect evidence to prohibit websites hosting such illicit material demonstrates a concerted effort to address the challenges posed by deepfakes and other forms of digital exploitation.

However, combating these crimes remains a formidable challenge for both the administration and

[15] Proposed Digital India Act 2023, Digital Indira Dialogues, 9th March,2023, MINISTRTY OF ELECTRONICS AND INFORMATION TECHNOLOGY, GOVERNMENT OF INDIA, https://www.meity.gov.in/writereaddata/files/DIA_Presentation%2009.03.2023%20Final.pdf

[16] In Sexual Violence Videos and Recommendations, (CRL) No (s). 3/2015 PRAJWALA Letter dated 18.2.2015.

the judiciary. Deepfake-related offenses are not only convenient and transnational but also elusive and difficult to trace. The lack of well-defined legal frameworks and precise definitions exacerbates the complexity of addressing these emerging threats. Moreover, the lucrative nature of digital crimes incentivizes perpetrators to exploit existing loopholes and evade detection and prosecution.

Despite these challenges, there is a growing global awareness of the perils posed by deepfake technology, prompting concerted efforts by corporations, research institutions, and law enforcement agencies to develop strategies for mitigating its impact. While significant strides have been made in combating online illicit content, the evolving nature of technical crime necessitates continuous reevaluation of forensic analytical methodologies, standards of evidence, and legislative mechanisms. Only through collaborative efforts between government, law enforcement, and tech stakeholders can we hope to effectively confront the expanding threat of deepfakeenabled criminality against women and society at large.

- ***Recent developments***[17]

The proliferation of deepfake content and other harmful AI media on social media platforms has prompted the Indian government to take decisive action to address the emerging threat. Minister of Electronics, Information, and Technology, Ashwani Vaishnaw, announced plans to implement a legal regulation aimed at detecting and limiting the spread of deepfakes and misleading synthetic content. This move comes in response to a surge in deepfake incidents targeting prominent personalities like actor Katrina Kaif and Rashmika Mandana, highlighting the urgency of penalizing creators of misleading synthetic content.

Tech-law specialist Khusbu Jain outlined the four pillars of the proposed regulation, emphasizing the importance of detection, prevention, reporting mechanisms, and awareness. Further, she highlighted the potential risks associated with deepfakes, including unauthorized access and compromising defense systems, underscoring the critical need for regulatory intervention to mitigate these risks.

[17] India takes action against deepfakes with new law, Dhaka Tribune (23rd November, 2023) India takes action against deepfakes with new law (dhakatribune.com) (accessed on 15.04.2024)

Minister Vaishnaw assured that the new regulation would undergo public consultation and be based on technology for effective detection. He emphasized the collaboration with social media platforms to detect deepfakes through technology, reflecting a coordinated effort to combat the misuse of synthetic media. Entrepreneur Alankar Saxena emphasized the broad impact of deepfakes on various industries, emphasizing the vulnerability of finance, technology, and media sectors. Saxena stressed the importance of advanced detection mechanisms, robust cybersecurity infrastructure, and awareness among employees to effectively combat the misuse of deepfake technology. AI expert Jaspreet Bindra warned about the potential misuse of deepfakes in upcoming elections worldwide, highlighting examples of deepfakes targeting political leaders. Bindra emphasized the need for governments to address the threat to democracy posed by manipulated content.

The Delhi High Court expressed skepticism regarding its authority to issue directives to regulate the use of deepfakes, suggesting that the government is better positioned to address the issue with a balanced approach. Acting Chief Justice Manmohan and Justice Mini Pushkarna, presiding over a Public Interest Litigation (PIL) petition seeking to block access to deepfake-generating websites, deliberated on the matter during proceedings. Acting Chief Justice Manmohan observed, "Given the ubiquity of this technology in the borderless realm, controlling the internet poses significant challenges. Policing it extensively risks encroaching upon the freedom of the internet. Therefore, there exist crucial, balancing considerations at play." Recognizing the government's awareness of the issue, the Court deferred further deliberation on the matter.

As the Indian government takes proactive steps to address the deepfake menace, the new regulation is expected to play a crucial role in safeguarding the integrity of digital content and protecting individuals from the malicious use of synthetic media.

6. UNITED STATES APPROACH TOWARDS LAWS RELATED TO DEEPFAKES

In recent years, numerous states across the United States have enacted legislation to regulate the potentially harmful use of deepfakes. However, the protection of free speech guaranteed under the

First Amendment poses significant constraints on the implementation of such laws[18]. For example, California's Assembly Bill No. 730[19], passed in 2019, made it illegal to use deepfakes in election materials within sixty days of the election, unless accompanied by a clear indication of their inauthenticity. While the law exempts media constituting satire or parody, it imposes penalties on those who distribute materially false campaign materials. Similarly, California granted individuals depicted in sexually explicit deepfake materials the right to pursue civil damages, regardless of their involvement in creating the content.

Virginia also revised its laws in response to deepfake threats, criminalizing the dissemination of altered photos with the intent to portray real individuals. New York has enacted legislation to protect individuals' digital likenesses under its Right to Privacy statute, which includes provisions for monetary compensation and injunctions against the unauthorized use of one's identity. Additionally, the United States Congress has proposed bills like the Identifying Outputs of Generative Adversarial Networks (IOGAN) Act and the Deepfakes Accountability Act to address the challenges posed by deepfakes. These bills aim to fund research on deepfake technology, mandate disclosures for deepfake creators, and impose civil penalties for non-compliance.

However, existing laws governing deepfakes in the United States have limitations, hindering victims' ability to seek relief. Legal recourse often requires victims to navigate complex jurisdictional issues and identify perpetrators, which can be challenging. Moreover, laws like the Communications Decency Act of 1996 limit liability for online platforms hosting deepfake content, further complicating enforcement efforts. Any legislative measures must carefully balance the need to combat deepfake threats with the protection of free speech rights guaranteed by the First Amendment. These laws should target instances involving malice or reckless disregard for truth and exclude situations deemed newsworthy to avoid infringing on constitutional rights.

---

[18] Penelope Thornton, "Deepfakes: An EU and U.S. perspective", GLOBAL MEDIA TECHNOLOGY AND COMMUNICATIONS QUATERLY (GMTCQ) SPRING SUMMER 2020, https://f.datasrvr.com/fr1/320/16758/1207330_-_GMCQ_-_Spring_2020_Deepfakes.pdf, pp 30-36.

[19] Kari Paul, "California Makes 'Deepfake' Videos Illegal, But Law May Be Hard to Enforce", The Guardian (Oct. 7, 2019), https://www.theguardian.com/us-news/2019/oct/07/california makes-deepfake-videos-illegal-but-lawmay-be-hard-to-enforce (accessed on 15.04.2024).

7. CONCLUSION AND SUGGESTION

It is well evident that the existing laws are not sufficient and hence with the increasing numbers in crime rate there is an urgent need of regulations to combat and criminalise the use of AI technology like creation of deepfakes. There should be stringent punishment rendered on deepfake pornography other related crimes done to hamper the privacy and which outrages the modesty of a women because the mostly the use of this technology has been done to hamper the women by objectifying them by creating fake images. The law makers are working towards an amicable solution but it must be robust as there's need of tools to identify deepfakes and even if the tool is prevalent technology is free, internet is free it will be very easy for an wrongdoer to get an antidote of that particular technology, hence while drafting laws all these consideration should be there in one's mind.

The emergence of deepfake technology presents a significant challenge in the realm of cybercrime, particularly against women, where it has been exploited for malicious purposes such as revenge porn, slut-shaming, sextortion and pornographic contents. The ease with which cybercriminals can manipulate videos and images using deepfakes poses a formidable technological battle for law enforcement agencies. The Indian legal framework currently lacks sufficient regulations to address the complexities of AI algorithms, leading to a rise in the misuse of deepfake technology. To combat this issue, several recommendations can be considered.

Firstly, governments should prioritize the implementation of anti-fake technologies to detect and combat deepfakes effectively. This may involve developing reliable methods for identifying deepfakes, such as AI-powered detection software or watermarking information to identify tampering. Additionally, training and awareness programs should be implemented to educate the public about the dangers of deepfakes and how to spot them. Collaborative efforts involving media, civil society, and government are essential to ensure widespread awareness and understanding of the risks associated with deepfake technology.

Moreover, leveraging blockchain technology could provide a viable solution for authenticating and verifying the legitimacy of digital media files. Blockchain's decentralized nature offers increased security and transparency, making it an effective tool for combating deepfake

manipulation. By utilizing blockchain to authenticate videos and audio files, individuals can verify the authenticity of media content and mitigate the spread of deepfake misinformation.

Furthermore, adopting a zero-trust stance towards internet content is crucial in combating the spread of deepfakes. Individuals should exercise extreme vigilance and skepticism when encountering online narratives, photos, and videos, considering the possibility of deepfake manipulation. Implementing identity checks and behavioral analytics can also help individuals detect and respond to deepfake threats effectively.

In conclusion, addressing the challenges posed by deepfake technology requires a multifaceted approach involving technological solutions, awareness programs, and collaboration between various stakeholders. By implementing these recommendations, governments can strengthen their efforts to combat cybercrime against women and protect individuals from the malicious use of deepfake technology.

REFERENCE:

A) STATUTES-

The Indian Constitution, 1951

Indian Penal Code, 1872

Information and Technology Act, 2000

Communications Decency Act of 1996

B) OTHER SOURCES

*Deepfakes-A-Challenge-for-Women-Security-and-Privacy-Dr.-Zubair-Ahmed-Khan-_-Ms.Asma-Rizvi[1].pdf

Articulating A Regulatory Approach to Deepfake Pornography in India (ijlt.in)

India takes action against deepfakes with new law (dhakatribune.com)

Laffier+&+Rehman+(2023)+-+Deepfakes+&+Harm+to+Women (1).pdf

Emerging Technologies and Law: Legal Status of Tackling Crimes Relating to Deepfakes in India |SCC Times (scconline.com) https://timesofindia.indiatimes.com/city/delhi/reel-is-real-dmrc-claims-clip-deepfake-2-womensay-otherwise/articleshow/108831003.cms

https://www5.austlii.edu.au/au/journals/UNSWLawJlStuS/2022/25.html