Deepfakes and their Ethical Implications: A Study among Youth in Mangalore

Dr Vidya Vinutha Dsouza[1], Mr. Glavin Dsouza[2]
[1]*Asst Professor, St Aloysius (Deemed to be University), Mangalore*
[2]*2nd MSW Student, St Aloysius College (Autonomous), Mangalore*

***Abstract*—Deepfake is a phenomenon where AI manipulates digital content to create realistic but false images, videos, and audio. This technology is increasingly influencing youth by spreading misinformation, enabling cyberbullying, and distorting reality. With social media as their primary information source, young people are highly vulnerable to deception and privacy risks. Chesney and Citron (2019) highlight the significant privacy and ethical concerns posed by deepfakes, particularly for youth, emphasizing the need for stronger legal protections. Paris and Donovan (2019) stress the vulnerability of young people to deepfake-generated disinformation due to a lack of media literacy. Similarly, Maras and Alexandrou (2019) discuss the psychological and reputational harm caused by non-consensual deepfakes, particularly among social media-active youth.**

The study employs a descriptive research design to understand the impact of deepfakes on the youth of Mangalore. A Stratified Random Sampling method was used, with 50 respondents selected to ensure representation across age groups, educational backgrounds, and urban/rural demographics. Data was collected through researcher-prepared questionnaires, complemented by secondary data from books and research articles.

The study highlights that 74% of respondents see deepfakes as a major threat to personal privacy, while 52% believe they severely harm individuals' reputations. Additionally, 66% recognize deepfakes as a contributor to online harassment among youth, raising concerns about their role in cyber abuse. To tackle the challenges of deepfakes, it is vital to promote digital literacy through workshops and awareness campaigns, equipping youth to identify manipulated content. Stricter regulations and advanced AI tools should be implemented to detect and flag deepfakes, ensuring a safer online environment.

***Index Terms*—AI Regulations, Cyber Harassment, Deepfake Awareness, Deepfake Ethics, Digital Literacy, Media Manipulation, Misinformation, Online Privacy, Social Impact, Youth Vulnerability.**

I. INTRODUCTION

The rise of artificial intelligence (AI) has revolutionized multiple domains, with deepfake technology emerging as one of the most controversial advancements. The term "deepfake" combines "deep learning" and "fake," referring to AI-generated visual and auditory content that mimics real-life images, videos, and voices with high precision. This technology, driven by generative adversarial networks (GANs), allows the digital manipulation of human likenesses, raising ethical concerns about misinformation, privacy, and public trust.

Initially perceived as a tool for entertainment and digital creativity, deepfakes have quickly become a societal and security challenge. The ability to fabricate realistic videos of individuals saying or doing things they never did has contributed to misinformation, cyberbullying, and reputational harm. Social media platforms amplify these risks, making deepfake content easily shareable and widely accessible.

A crucial demographic affected by deepfakes is youth, particularly those in digitally connected regions such as Mangalore. As a growing urban centre in India, Mangalore has seen a rise in internet penetration and smartphone usage, exposing its young population to both the opportunities and threats of deepfake technology. With platforms like Instagram, YouTube, and TikTok shaping youth culture, the increasing presence of deepfakes poses risks to their ethical values, digital literacy, and emotional well-being.

This research examines the ethical implications of deepfakes, focusing on misinformation, privacy violations, cyberbullying, and trust erosion among Mangalore's youth. By exploring legal frameworks, social media policies, and educational interventions, this study seeks to address the need for media literacy and regulatory measures to mitigate the risks of

deepfake technology. Understanding how young people engage with and perceive deepfakes is crucial to developing strategies that protect them from the ethical and psychological challenges posed by this evolving digital phenomenon.

II. REVIEW OF LITERATURE

Deepfake technology poses significant ethical and psychological concerns, particularly for young people. Donath (2019), in *The Ethics of Digital Deception*, argues that deepfakes create "shadow identities" that distort personal authenticity. He highlights the dangers of cyberbullying and harassment, emphasizing the need for stronger ethical laws to protect vulnerable youth, particularly in Mangalore, where digital engagement is high.

Whittaker et al. (2020), in *Deepfakes and the Future of Fake News*, explore how deepfake- driven misinformation affects young people, the largest consumers of online news. They stress that deepfakes undermine public trust in media and democratic institutions. To counteract this, they advocate for media literacy programs that enable youth to critically analyze digital content, a crucial need for Mangalorean youth navigating social media landscapes.

Suler (2004), in *The Online Disinhibition Effect*, examines how digital interactions lower self-regulation, making young people more susceptible to deepfake-related harms. His research suggests that deepfakes amplify online disinhibition, leading to the misuse of manipulated media for deception and harassment. He recommends incorporating ethical discussions on digital responsibility into youth education, particularly in Mangalore, to prevent deepfake misuse.

III. OBJECTIVES OF THE STUDY

1. To understand the demographic profile of the respondents
2. Assess youth familiarity with and recognition of deepfakes in Mangalore
3. Explore youth perceptions of deepfake ethics, including privacy and consent
4. Investigate psychological and social impacts of deepfakes on youth
5. Evaluate youth awareness of deepfake laws and protection measures

IV. METHODOLOGY

The study employs a descriptive research design to understand the ethical and psychological challenges posed by deepfakes among Mangalore's youth. The study focuses on youth residing in Mangalore. The estimated youth population is 136,183 (Census 2025). Using Convenient sampling, 50 respondents were selected to ensure diverse representation.

Data was collected through self-prepared questionnaire and secondary data through Research articles, Journals and online sources

Major Findings

Table 4.1 Demographic Profile

Demographic data provides key insights into the study population and helps contextualize research findings. This study examines youth in Mangalore, focusing on age, gender,
education, residence, and social media usage to understand their exposure to deepfakes and ethical perceptions.

The age distribution is balanced, with 30% aged 15–18, 40% aged 19–21, and 30% aged 22– 24, ensuring diverse perspectives. The largest group, 19–21 years (40%), is actively engaged with digital platforms, making them more likely to encounter deepfakes.

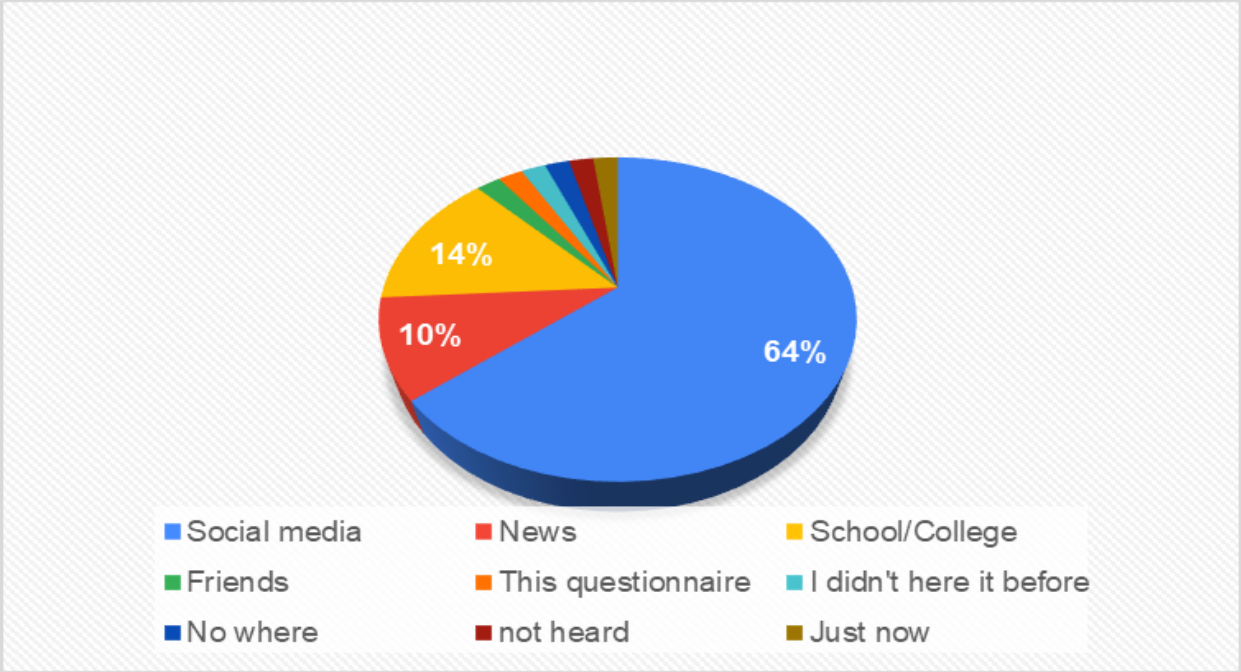
Gender

representation includes 58% female, 40% male, and 2% undisclosed, reflecting a higher female participation in discussions on digital ethics.

Education-wise, 46% are undergraduates, 28% postgraduates, and 26% pre-university
students, highlighting varying levels of digital awareness. 56% reside in urban Mangalore and 44% in rural areas, ensuring diverse viewpoints. With 80% using social media daily, the study underscores the need for stronger digital literacy initiatives to address ethical concerns
surrounding deepfake technology.

Variables	Classification	Frequency	Percent
Age(N50)	15–18	15	30%
	19–21	20	40%
	22–24	15	30%
Gender(N50)	Male	20	40%
	Female	29	58%
	Prefer Not to Say	1	2%
Education(N50)	Pre-University	13	26%
	undergraduate	23	46%
	postgraduate	14	28%
Residents(N50)	Urban Mangalore	28	56%
	Rural Mangalore	22	44%
Social Media Usage(N50)	Daily	40	80%
	Rarely	3	6%
	Weekly	7	14%

Diagram 4.2.1 Source of Exposure to Deepfakes



The rise of deepfake technology has raised concerns about misinformation, digital ethics, and media manipulation. Chesney and Citron (2019) argue that deepfakes undermine digital trust by making it difficult to distinguish between real and AI-generated content. Understanding how youth in Mangalore perceive deepfakes is crucial in assessing their awareness, ethical concerns, and the need for digital literacy.

The findings indicate that 56% of respondents have heard of deepfakes, while 44% have not, highlighting a significant knowledge gap. While 62% correctly identified deepfakes as AI- generated fake media, only 12% recognized their potential for misinformation, showing limited awareness of ethical risks. Additionally, 26% of respondents lacked a clear understanding of deepfakes, suggesting the need for better education. Given their increasing prevalence, schools, media campaigns, and digital literacy programs should focus on educating youth about deepfake detection and responsible engagement with AI-generated content.

Table 4.3.1 Deepfakes: Privacy Threats vs. Ethical Concerns

Variables	Classification	Frequency	Percent
Deepfakes pose a threat to personal privacy	Yes	6	12%
	No	28	56%
	Not Sure	16	32%
How ethical it is to create deepfake videos for entertainment purposes	Ethical	37	74%
	Unethical	3	6%
	Neutral	10	20%

Findings reveal that 56% of respondents do not see deepfakes as a privacy risk, 32% are unsure, and only 12% acknowledge the threat, highlighting a lack of awareness about AI-generated identity theft and misuse. Additionally, 74% consider deepfakes ethical in entertainment, suggesting youth associate them with humor and visual effects rather than ethical concerns. However, 6% view deepfake entertainment as unethical, recognizing risks of misinformation and consent violation.

Table 4.5.1 Confidence in Legal Protection against Deepfakes Across Age Groups

How confident are you in current laws to protect individuals from deepfakes?	Age			Total
Classification	15–18	19–21	22–24	
Very confident	0	4	0	4
Somewhat confident	8	10	11	29
Not confident at all	7	6	4	17
Total	15	20	15	50

V. SUGGESTIONS

A study among youth in Mangalore reveals low confidence in current legal protections against deepfakes, with only 8% feeling "very confident." The majority (58%) are "somewhat confident," while 34% are "not confident at all." Confidence varies by age, with 20% of 19– 21-year-olds expressing high confidence, while younger groups (15–18 and 22–24) show scepticism. The 15–18 group is particularly uncertain, with 47% lacking confidence, possibly due to limited awareness of cyber laws. The findings highlight a general lack of trust in existing regulations and a need for stronger legal frameworks to address deepfake misuse. To improve confidence, digital law awareness should be incorporated into education, and clearer, stricter regulations must be established. Policymakers, legal institutions, and tech companies should collaborate to enhance legal protections and ensure youth are safeguarded against AI-generated misinformation and manipulation.

- Develop media literacy programs for fact-checking and misinformation detection – Encouraging youth to question and verify the authenticity of videos and images will help prevent the spread of false information.
- Strengthen reporting mechanisms for deepfake misuse – Youth should have accessible and anonymous options to report harmful deepfake content, with social media platforms responding promptly to remove or verify it.
- Develop global cooperation for deepfake regulation – Since deepfake content can spread beyond national borders, governments should collaborate internationally to establish common legal frameworks for tackling AI-generated misinformation.
- Introduce age-restricted content filtering for deepfakes – Since youth are more vulnerable to misinformation, platforms should enable content restrictions that filter out harmful AI-generated

media.

- Foster collaboration between AI researchers, ethicists, and legal professionals – A multidisciplinary approach should be adopted to ensure AI advancements are made responsibly while minimizing ethical risks.

VI. CONCLUSION

This study on Deepfakes and Their Ethical Implications: A Study Among Youth in Mangalore highlights the growing risks of AI-generated content, including misinformation, privacy violations, and ethical concerns. While youth are increasingly exposed to deepfakes, many lack awareness of their risks and legal consequences. The findings emphasize the need for stronger media literacy programs, improved social media regulations, and stricter legal frameworks to prevent deepfake misuse. By promoting responsible AI use, enhancing digital education, and enforcing ethical guidelines, society can ensure that deepfakes are used ethically, empowering youth to critically engage with digital content while minimizing potential harm.

REFERENCE

[1] Chesney, R., & Citron, D. K. (2019). Deepfakes and the new disinformation war: The coming age of post-truth geopolitics. *Foreign Affairs*, 98(1), 147-155. https://www.foreignaffairs.com/articles/world/2018-12-11/deepfakes-and-new-disinformation-war

[2] Deep trace Labs. (2019). *The state of deepfakes: Landscape and impact*. https://sensity.ai/reports/

[3] Donath, J. (2019). The ethics of digital deception: Deepfakes and the online manipulation of identity. *AI & Society*, 34, 679-687.https://link.springer.com/article/10.1007/s00146-019-00888-1

[4] Dsouza, V. V., & Varghese, A. (2022). Impact of Online Class during Covid-19 Lockdown. *International Journal of Early Childhood Special Education*, *14*(4).

[5] Hinduja, S., & Patchin, J. W. (2010). Bullying, cyberbullying, and suicide. *Archives of Suicide Research*, 14(3), 206–221. https://doi.org/10.1080/13811118.2010.494133

[6] Jiang, F., Shen, Y., & Chen, Y. (2020). Deepfakes and AI ethics: Implications for personal security. *Journal of Cybersecurity*, 8(2), 109-https://academic.oup.com/cybersecurity/article/6/2/tyaa021/5932311

[7] Khatoon, S., & Dsouza, V. V. (2018). Economic & Health consequences of donating a Kidney: A qualitative study. *International Journal of Research in SocialSciences*, *8*(4), 617-621.

[8] Lupton, D., & Williamson, B. (2017). The datafied child: The digital transformation of childhood. *Technology, Knowledge and Learning*, 22(2), 139-155.https://link.springer.com/article/10.1007/s10758-017-9320-0

[9] Pallavi, A., & D'Souza, V. V. (2018). Inclusive workplace-A fair shot for disabled. *International Journal of Research in Social Sciences*, *8*(2), 411-419.

[10] Paris, B., & Donovan, J. (2019). Deepfakes and cheap fakes: The manipulation of audio and visual evidence. *Data & Society Research Institute*. https://datasociety.net/library/deepfakes-and-cheap-fakes/

[11] Suler, J. (2004). The online disinhibition effect. *Cyber Psychology & Behavior*, 7(3), 321-https://www.liebertpub.com/doi/abs/10.1089/1094931041291295

[12] Whittaker, J., O'Loughlin, B., & Gillespie, T. (2020). Deepfakes and the future of fake news: Navigating the digital disinformation landscape. *Digital Journalism*, 8(3), 379-https://www.tandfonline.com/doi/full/10.1080/21670811.2019.1691921