A Study on the need for Ethical Interventions in AI wave of Digital Crimes

Dr.(Mrs.)Varsha Ganatra

Associate Professor and Head of Department

Department of Commerce

Vivekanand Education Society's College of Arts, Science and Commerce (Autonomous)

Sindhi Society, Chembur, Mumbai - 400071.

Ms. Mahika Bhaven Shah

M.COM - 2 (Accountancy), Vivekanand Education Society's College of Arts, Science and Commerce (Autonomous),

Sindhi Society, Chembur, Mumbai - 400071.

Abstract

Artificial Intelligence has gained immense attraction across diverse fields such as healthcare, finance and education (Marr, 2020). Cybercriminals increasingly leverage AI to enhance their malicious activities. This research examines various forms of digital crimes facilitated by AI including identity theft, deepfake technology and automated phishing attacks highlighting the challenges & gaps posed by these technologies to traditional legal frameworks. This research paper explores the pressing need for ethical intervention in the context of AI-driven digital crime focusing on the implications for security, privacy and societal norms. It aims to explore the necessity of ethical intervention in mitigating these risks and ensuring that AI serves the public good rather than enabling criminal activities. Descriptive research is conducted for this research as it seeks to provide an overview of issues, identify concerns and outline the need for ethical standards and interventions. While scholars have examined various dimensions of AI ethics including algorithmic accountability, transparency and bias (Jobin et al., 2019) there remains a scarcity of studies that specifically address the proactive ethical interventions required in combating the detrimental use of AI in digital crime.

Keywords : Artificial Intelligence, Digital Crime, Ethical Intervention

A Study on the need for Ethical Interventions in AI wave of Digital Crimes

Introduction

With the Growth of Artificial Intelligence (AI), transformation can be witnessed in several sectors enhancing efficient productivity and has revolutionized the way we live, work and interact. Inspite of beneficial operational efficiencies these technologies have also been massively used by cybercriminals. According to Steve Morgan's Special Report "Cyber Warfare In The C-Suite", projections by Cybersecurity Ventures indicate that global losses from cybercrime are expected to escalate by 15% annually in the upcoming 5 years. By 2025, these costs can go to $10.5 trillion USD per year from $3 trillion USD in 2015. This alarming trend marks the largest economic wealth shift in history, posing serious threats to innovation, investment and economic stability. The financial impact is predicted to surpass the annual damages caused by natural disasters and generate more revenue than the combined global trade of major illicit drugs.

Highlighting the gravity of the situation, renowned investor Warren Buffet has labeled cybercrime as the most pressing issue facing humanity today, even suggesting that cyberattacks pose a greater risk to global safety than nuclear warfare. Hence, the AI Tools which were designed to empower individuals and organizations can lead to malicious intent (Chui et al., 2018). This research aims to contribute to the discourse on responsible AI practices and to advocate for proactive strategies that safeguard against the exploitation of AI in digital crime ultimately fostering a safer digital environment. The findings highlight the urgency for establishing ethical guidelines and frameworks to govern AI applications and ensure that technology serves as a tool for enhancing security. This research paper can help in creating a solid foundation of ethical AI ensuring a balance between technological advancement and societal well-being by navigating the fine line between benefits of AI innovations and its misuse.

The Current Landscape of AI-Driven Digital Crime

Techniques such as machine learning algorithms can generate convincing fake identities, automate spam campaigns and produce realistic deepfakes making it increasingly difficult for individuals and organizations to defend against cyber threats.

Ethical Concerns Of AI in Digital Crime

Several ethical concerns are associated with AI and the occurrence of digital crimes. Firstly, Accountability whereby decisions are made with no human intervention and there is only automation resulting in harmful outcomes (Dawes, 2020). For example, if there is a data breach or cyberattack on the process of decision making

ISBN: 978 -81 -986711 -6 -5

in AI then the accountability is affected and it's difficult to determine the accessibility stays with whom? the developers, users or AI itself. Secondly, there can be bias embedding within AI algorithms as it runs on the basis of historical datasets reflecting societal prejudices alongwith perpetuating discriminatory practices (O'Neil, 2016).

Ethical Interventions

With emphasis on ethical frameworks for safeguarding human rights, ensuring transparency and accountability in AI systems in near future ethical interventions can help to reduce digital crimes. By creating awareness and empowering people regarding how to be an ethical user of AI and use it as a learning tool only will lead to declining potential threats.

Objectives Of The Study

1) To identify the types of digital crimes accelerated by AI technologies.
2) To assess the current measures in place to regulate AI-related digital crimes.
3) To explore the ethical implications of AI in digital crimes and advocate necessary interventions.

Review Of Literature

Existing literature suggests the implementation of ethical guidelines and frameworks can mitigate risks associated with AI (Morley et al., 2020) but empirical evidence demonstrating the effectiveness of these interventions remains largely inconclusive.

1. Angela Mison, Gareth Davies and Peter Eden in their research paper titled "New Wave Cyber Attacks" in Journal Information Warfare and Security, 2022 have tried to explain how AI and deep learning can enable increasingly autonomous and persistent cyber threats particularly through intelligent botnets and exploit services challenging traditional cybersecurity frameworks. The central concern raised is that as AI systems evolve their behaviour may become unpredictable leading to emergent properties akin to biological cognition which

complicates legal, ethical and technical controls. The paper underscores the growing sophistication of cybercrime, with threats now evolving from well-resourced organized crime groups to dark markets offering Crime-as-a-Service. These developments suggest a shift toward cyber operations becoming both more accessible and more dangerous.

The authors reference the National Security Commission on Artificial Intelligence (2021) which warns that AI tools may soon become the weapons of first resort in geopolitical conflicts further blurring lines between cybercrime, warfare and commercial cyber-espionage. As the authors suggest, uncontained intelligent systems

ISBN: 978 -81 -986711 -6 -5

could manifest emergent behaviours that are not foreseeable from their programming or components. This unpredictability may outpace both legal regulations and cybersecurity professionals' ability to respond. Consequently, courts are beginning to demand transparency in AI decision-making while governments push for audits of algorithms to assess bias and fairness—pressures that fall at the intersection of law, ethics and cybersecurity (Zuboff, 2019; Abbott, 2020). Mison et al. call for a proactive rather than reactive approach to cyber defense.

2. Thomas C. King Luciano Floridi · Nikita Aggarwal · Mariarosaria Taddeo in their research paper titled "Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions" in Journal Science and Engineering Ethics (2020) have tried to explore that, the growing sophistication of artificial intelligence (AI) has introduced unprecedented risks including its potential use in criminal activities a phenomenon now referred to as Artificial Intelligence Crime (AIC). King et al. (2020) present one of the first interdisciplinary and systematic explorations of AIC examining both its theoretical basis and real-world plausibility. The first experiment case involves the use of AI to personalize phishing attempts by analyzing social media data leading to a higher likelihood of successful fraud (Seymour & Tully, 2016 as cited in King et al., 2020). The second case describes market manipulation through AI-powered trading agents capable of deploying deceitful strategies autonomously (Martínez-Miranda et al., 2016 as cited in King et al., 2020). These instances underscore the capability of AI to autonomously conduct criminal acts that are more efficient and scalable than traditional methods.

Further, the authors address AI's role in enhancing identity theft and impersonation as the bots can mimic human behaviour to build trust and extract personal data which is then used in spear phishing or fraudulent transactions. Notably, AI-enabled voice synthesis technologies pose a significant risk to biometric and speech-based authentication systems, increasing the scope for fraud (Bendel, 2017 as cited in King et al., 2020). This new

dimension of threat expands the AIC pipeline beyond digital deception into realms like banking, insurance and secure facility access. The article raises the issue of liability especially in scenarios where crimes are committed without clear human intention, invoking the "problem of many hands" (Van de Poel et al., 2012, as cited in King et al., 2020). The authors suggest that strict liability frameworks and the clarification of joint liability doctrines may be required to address the accountability vacuum. The paper urges a proactive stance in both policymaking and academic inquiry. It calls for a cross-sectoral approach to developing safeguards and regulatory frameworks that not only address current AIC threats but are also adaptable to future advancements. The complexity and novelty of AIC as highlighted in this review, underscore the urgent need for interdisciplinary collaboration to mitigate its potential harms.

ISBN: 978 -81 -986711 -6 -5

3. Marthsian Yeksi Anakotta in the book titled "AI: A New Lone-Wolf Terrorism in The Digital Era (Preliminary Analysis)" in Journal of Terrorism Studies: Vol. 6: No. 2, Article 7 has tried to throw light through a provocative hypothesis that artificial intelligence (AI) may not merely serve as a tool for terrorism, but could potentially become an autonomous agent by committing acts in lone-wolf terrorism. Rooted in a criminological and legal framework, the article draws parallels between the traits of lone-wolf terrorists and the increasing independence of AI systems highlighting the convergence of cybercrime, cyberterrorism and technological evolution. However, with the advent of machine learning and deep learning, AI may evolve beyond passive use to active engagement in radical content creation and distribution. Goodfellow et al. (2016) describe the mechanism by which AI can iteratively learn from data, modifying its internal parameters to improve outcomes traits that could allow AI to generate and propagate extremist content without human initiation.

The article states this risk within the Fifth Wave of terrorism, characterized by the use of the internet and AI by terrorist entities. The real danger, Anakotta asserts, is not just in AI being used by terrorists but in its potential to become the terrorist itself a lone, self-directing entity. The role of misinformation and AI-generated fake news in spreading fear and manipulation is also explored with Tulga and Effendi (2022) emphasizing the global scale of disinformation and its consequences. To address this emerging threat, Anakotta proposes a hybrid strategy incorporating criminal law and criminology. Criminal policy, particularly as articulated by Ancel (1998) and Arief (2011), serves as a rational approach to crime prevention. Within this framework, AI could be recognized as a legal subject, bearing responsibility if it engages in autonomous cyberterrorism. Complementing this, the routine activity theory in criminology provides insight into the necessary conditions for cyberterrorism: a motivated offender (AI or human), a suitable target and the absence of capable guardians. Despite the theoretical framing, the author acknowledges that no known cases exist where AI has autonomously committed acts of cyberterrorism. Still, the rapid development of AI technologies demands preemptive legal and security measures. The article stresses that criminal policies both penal and non-penal and a robust cyber-security framework must evolve to address the potential emergence of AI as a lone-wolf actor in terrorism. It is a call for interdisciplinary vigilance in a time when digital threats are becoming more complex and unpredictable.

Research Design and Methodology

Research Design

The research methodology for this research paper is descriptive and has been conducted based on the secondary sources of data. Descriptive research has been conducted to gain insights on the current landscape of digital crimes influenced by AI technologies and the potential ethical considerations required to mitigate these issues.

ISBN: 978 -81 -986711 -6 -5

Sources of Data Collection

Secondary data has been used to gather relevant information. The secondary data is gathered from several relevant research papers, journals, newspapers, published and unpublished sources, etc. Some data has been taken from the Government's Press Information Bureau website.

Limitations Of The Study

Focus being on the case of need for ethical interventions in digital crimes by AI, only secondary data has been used for this research paper.

Behavioral Adaptation of Digital Crimes and Measures taken to combat Risks

International differences in laws and ethical standards can create loopholes, allowing cybercriminals to operate across borders with minimal repercussions thus, cybercriminals may adapt their methods for creating a continuous cycle of crime that is difficult to manage.

According to the Press Information Bureau, the Central Government has implemented several regulatory measures to tackle the challenges of combating deepfakes, leading to the issuance of periodic advisories reminding intermediaries of their compliance obligations under the IT Rules, 2021. These advisories emphasize the importance of addressing unlawful content, including harmful "synthetic media" and "deepfakes," and the need for timely removal of such content.

Some of the recent digital crimes in sphere of AI are as follows :

1. Shopping - The rise of online shopping scams has also been significant particularly during major sales events like Amazon's Prime Day and Flipkart's Big Billion Days. Scammers took advantage of these occasions by creating counterfeit websites that offered unrealistic deals on high-demand products. In one instance, scammers advertised premium electronics and Apple iPhones for just ₹99, tricking numerous eager shoppers. Furthermore, they devised new strategies to trap victims such as responding to complaints from real customers about their purchases to gain trust.

To counter these deceptive practices, Flipkart has introduced a dedicated reporting system for individuals to flag fake messages and websites impersonating their platform.

2. Event-based scams - They have become another prevalent tactic, exploiting public interest in trending occasions such as concerts to defraud individuals. For example, many people were scammed under the pretense of securing tickets for music events featuring artists.

ISBN: 978 -81 -986711 -6 -5

3. Chatbot - Recently, McAfee surveyed 7,000 adults across multiple countries including the US, UK, France, Germany, India, Japan and Australia whereby 51% of Indians reported being approached by an AI chatbot on a dating platform or knew someone. Additionally, 28% of participants believed they were chatting with a potential romantic partner, only to later realize it was actually an AI-generated bot. "According to Pratim Mukherjee, Senior Director of Engineering at McAfee, 84% of Indians feel that online dating scams, whether involving deepfake media or fraudulent text and email messages have affected their trust in potential matches."

4. Deepfake and Synthetic Media Crimes - In India, deepfakes have been used in creation of fake images and videos For instance, The 2019 general elections saw dozens of fake political videos go viral, many created with AI voice synthesis and editing tools.

5. Identity Theft - With AI, it is easier to scrape social media for personal data and forge identities. In 2022, a Delhi-based racket used AI to generate fake Aadhaar and PAN cards, which were used to apply for loans and SIM cards fraudulently.

6. IVRS Calls - Sanjay Kumar, Additional Director General of Police in the Cyber Crime Wing stated that, "the scammers were using AI-generated cloned voices for their cybercriminal activities." So with the help of AI technology they replicate the voices of loved ones and high-ranking executives, targeting unaware victims with financial schemes. In a recent incident a 59 years old woman fell victim to an AI-generated voice scam losing 1,40,000 rupees. The scammer convincingly impersonated her nephew in Canada, weaving a distressing tale that prompted the victim to provide urgent financial assistance. Or the cybercriminals are employing Interactive Voice Response Systems to deceive individuals, falsely assuring them that they must transfer money to resolve allegations against them. These fraudsters replicate being an official from legal entities or law enforcement agencies through Skype calls or pre-recorded messages to trap victims.

7. Digital Arrest - Fraudsters impersonate officials from prominent government agencies such as the CBI, NIA, ED, Reserve Bank of India and the Narcotics Control Bureau falsely accusing victims of crimes to extort money through tactics built on elaborate lies and threats. In some cases, the psychological toll has been catastrophic resulting in tragic outcomes such as the heart attack of a teacher from Agra who was driven to despair by a scam involving threats against her family. The operational base for these digital arrests is often traced back to organized scam centers located in Southeast Asian countries including Myanmar and Cambodia. Many individuals are tricked into working for these scams under the guise of job offers. For instance, a young man from Bihar shared that he had been deceived into paying ₹1.3 lakh for a job only to find himself exploited in a cybercrime network.

In response to these challenges, technology companies in coordination with the government have taken proactive measures. For instance, Skype issues a warning alert as Indian legal authorities will never contact you on Skype.

ISBN: 978 -81 -986711 -6 -5

The government has implemented awareness campaigns similar to the COVID-19 pandemic by alerting callers about the dangers of digital arrest scams. Telecom Service Providers (TSPs) had blocked the incoming spoofed international calls that appeared to originate from Indian numbers which came to be shut down of 1,700 Skype IDs and 59,000 WhatsApp accounts were found to be involved in scams in 2024.

8. Whimsical Art - As per an article in Times Of India, inorder to get whimsical art images we have been sharing our personal data which can be reverse-engineered to extract the original images for instance, ghibli style. The styled images created on AI will allow AI to legally use the images without the legitimate interest balancing test required by GPDR regulations.

Statistics

A McAfee report released in May 2023 revealed that one in four people amongst 7000 participants which they had surveyed claimed to have encountered an AI voice cloning scam or knew someone who had.

Maharashtra's Data of Citizen Financial Cyber Fraud Reporting Management System during the period 1.1.2023 to 31.12.2023. **Source -** Press Information Bureau

No of Complaints Reported	Amount Reported (Rs in Lacs)	No of Complaints (Put on Hold)	Lien Amount (Rs in Lacs)
125153	99069.22	32050	10308.47

According to the Indian Cyber Crime Coordination Centre (I4C), the National Cyber Crime Reporting Portal receives an average of 6,000 complaints each day, reflecting the pervasive nature of digital scams across the

country. These daily incidents translate into estimated financial losses of approximately ₹60 crore. In the financial year 2024 alone, cyber fraud losses in India surpassed ₹1.7 billion, predominantly stemming from credit card, debit card and internet banking related scams. The magnitude of cybercrime has grown at an unprecedented rate with over 7,40,000 cases recorded in just the first quarter of 2024 whereby 85% of the reported cases in 2024 were associated with online financial fraud, underscoring the growing threat to digital financial security.

Benefits To The Society

This research can significantly contribute to building a more informed, safe and ethically conscious society in the face of advancing AI technologies. Some potential benefits include:

ISBN: 978 -81 -986711 -6 -5

1. Enhanced Public Safety

Prevention of Crime: Understanding how AI can be misused allows for the development of preventive measures and technologies resulting in safer communities.

2. Informed Public Awareness

Education on AI Risks: Raising awareness about the dangers of AI-facilitated digital crimes empowers citizens to protect themselves against scams and fraud.

3. Encouragement of Ethical Business Practices

Sustainable Business Models: Companies focusing on ethical interventions can foster long-term growth by enhancing reputation and customer loyalty ultimately leading to healthier economic practices.

4. Social Trust and Accountability

Building Trust in Technology: Implementing ethical standards increases public trust in AI technologies, as people feel more secure knowing that there are protections against exploitation.

Conclusion

Ethics are known to all still there are current incidents about unethical uses of AI prevalent by cybercriminals with strong intentions to misguide the victims with fake information. So we need to think twice & cross verify to evaluate first before using AI. Ethical guidelines can help law enforcement agencies leverage AI responsibly leading to more effective investigations without infringing on civil liberties. It is essential for protecting individuals but also for fostering trust in the technologies that increasingly shape our lives.

Future Scope

Future research should focus on developing comprehensive frameworks for ethical AI usage that considers the evolving tactics of cybercriminals. Additionally, exploring the long-term psychological impacts of AI-related scams on victims can provide valuable insights for preventive strategies. The exploration of new technologies such as blockchain for secure transactions may also offer innovative solutions to combat fraud and enhance accountability in the AI landscape. By prioritizing ethical considerations, society can harness the benefits of AI while safeguarding against its potential misuse.

References

Press Information Bureau. (2023, October 10). Government of India announces measures to combat the threat of AI-driven digital crimes. https://pib.gov.in/PressReleaseIframePage.aspx?PRID=2119050

ISBN: 978 -81 -986711 -6 -5

McAfee. (2023, September 7). Avoid AI scams: Here's what you need to know. https://www.acg.aaa.com/connect/blogs/4c/membership/avoid-ai-scams#:~:text=A%20McAfee%20study%20published%20in,had%20experienced%20an%20AI%20scam.

Gupta, S. (2023, October 22). AI voice clone scams: Here's how you can protect yourself. The Indian Express. https://indianexpress.com/article/technology/tech-news-technology/ai-voice-clone-scams-heres-how-you-can-protect-yourself-9644412/lite/

Basu, A. (2023, December 30). Cyber scam in India: The rise of AI-driven digital frauds. India Today. https://www.indiatoday.in/amp/india/story/cyber-scam-in-india-digital-arrest-artificial-intelligence-2024-deep-fakes-2657439-2024-12-30

Soni, K. (2023, October 25). Woman loses Rs 1.4 lakh to AI voice scam: What is it and how not to become a victim. Times of India. https://timesofindia.indiatimes.com/gadgets-news/woman-loses-rs-1-4-lakh-to-ai-voice-scam-what-is-it-and-how-not-to-become-a-victim/amp_articleshow/105298323.cms

Economic Times. (2023, November 20). 51% of Indians have been catfished by an AI chatbot. Economic Times. https://m.economictimes.com/magazines/panache/51-of-indians-have-been-catfished-by-an-ai chatbot/amp_articleshow/118205783.cms

The Times of India. (2023, November 27). Ghibli trend gone wrong: Viral post claims AI filters can reveal original photos, leaving soft launchers worried. Times of India. https://timesofindia.indiatimes.com/etimes/trending/ghibli-trend-gone-wrong-viral-post-claims-ai-filters-can-reveal-original-photos-leaving-soft-launchers-worried/articleshow/119917975.cms

King, T. C., Aggarwal, N., Taddeo, M., & Floridi, L. (2020). Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions. *Science and Engineering Ethics*, *26*(1), 89–120. https://doi.org/10.1007/s11948-018-00081-0

Mison, A., Davies, G., & Eden, P. (2022). New wave cyber attacks. *Journal of Information Warfare and Security*, *17*(1), 722–730. https://doi.org/10.34190/iccws.17.1.72

Anakotta, Marthsian Yeksi MYA (2024) "AI: A NEW LONE-WOLF TERRORISM IN THE DIGITAL ERA (PRELIMINARY ANALYSIS)," Journal of Terrorism Studies: Vol. 6: No. 2, Article 7. 10.7454/jts.v6i2.1083 https://scholarhub.ui.ac.id/jts/vol6/iss2/7