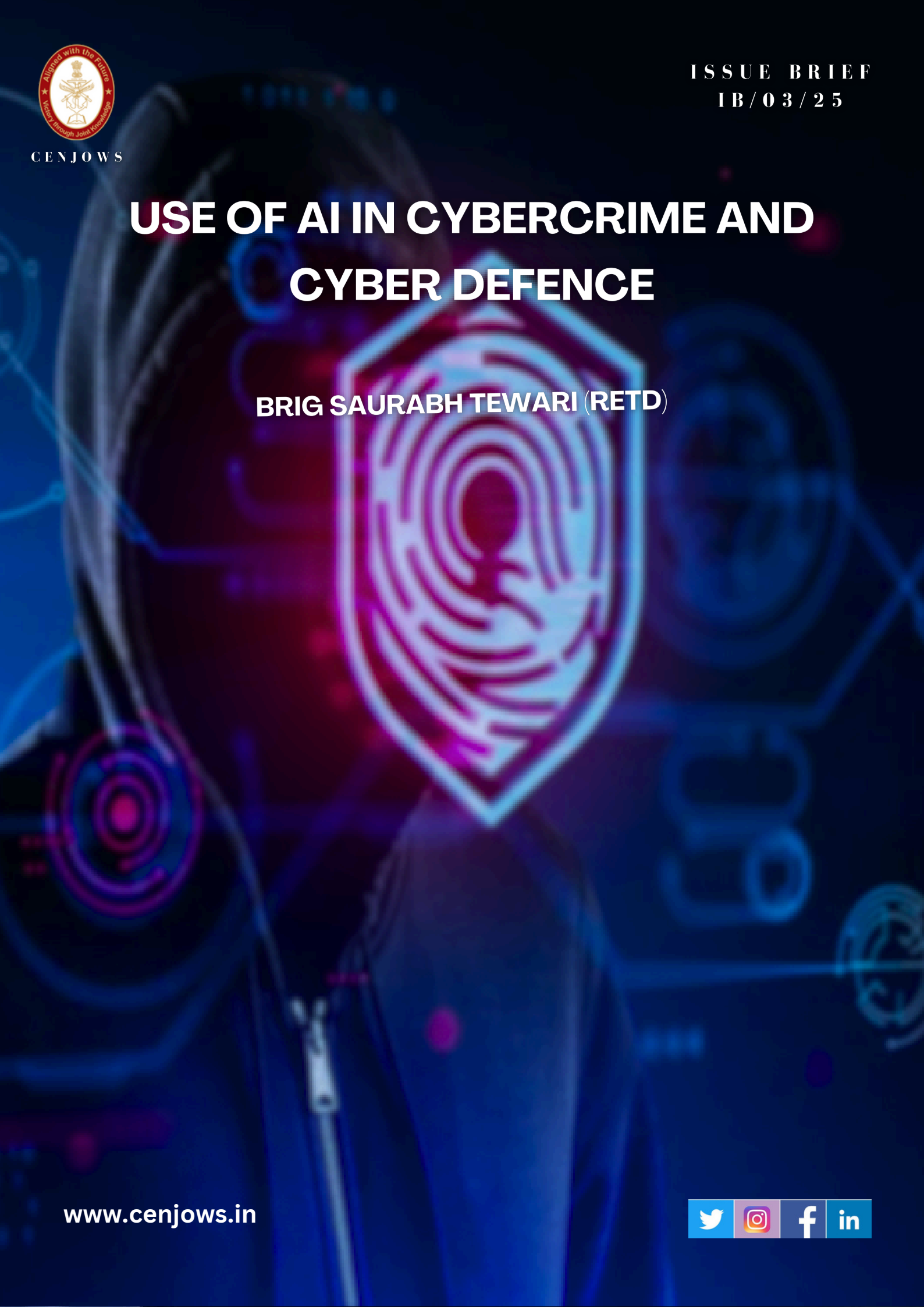
CENJOWS

ISSUE BRIEF
IB/03/25

USE OF AI IN CYBERCRIME AND CYBER DEFENCE

BRIG SAURABH TEWARI (RETD)



www.cenjows.in

CENTRE FOR JOINT WARFARE STUDIES

CENJOWS

USE OF AI IN CYBERCRIME AND CYBER DEFENCE

Brig Saurabh Tewari (Retd), is recognised for his contributions to military technology and cyber warfare, and is currently an Advisor at Army HQ.

Abstract

Artificial Intelligence (AI) is touted to be a breakthrough technology of this century that will change the way we think, act, conduct our daily business and chores and much more. Basic AI combined with deeper technologies like the Machine Learning (ML), Deep Learning, Large Language Model (LLM), Natural Language Processing (NLP) and Generative AI (Gen-AI) are going to make AI even more intelligent and human like, probably going beyond even what a human brain can do, and much faster and accurate. Creation of deep fake videos, cloned voice, polymorphic malware and intelligent phishing emails are just some examples that are already showcasing the dark side of AI. As AI becomes more mature, things will change drastically. However, AI can be put to good uses as much as it is being used by cyber criminals; it can be used to predict cyber-attacks, build a proactive and predictive defensive strategy, automate incident response, detect malware, analyse traffic patterns for anomalies, flag human behaviour anomalies for predicting insider threats, etc. This paper dwells into the use of AI for cybercrime & cyber-attacks and the good uses of AI for cyber security.

Keywords: *Adversarial Attacks, Artificial Intelligence (AI), Chatbots, ChatGPT, Deep Fake, Deep Learning, Distributed Denial of Service (DDoS) Attack, Generative Adversarial Network (GAN), Generative AI (Gen-AI), Generative Pre-trained Transformer (GPT), Graphical Processing Units (GPUs), Large Language Models (LLMs), Machine Learning (ML), Polymorphic Malware, Ransomware, Social Engineering, Tensor Processing Units (TPUs), Threat Intelligence.*

What is Artificial Intelligence (AI)?

Artificial intelligence (AI) is a new technology that empowers computing machines to train on huge data sets provided to them and cape human wisdom & knowledge, problem resolving, decision making, imagination and cognitive thinking. AI enabled devices, machines and applications can see and recognise objects, they can recognise and react to human voice. They are capable of learning from new incidents, knowledge, information and data. They can take independent decisions like human beings by absorbing real-time events, applying logic and reacting in real time. A common example is a self-driving car or voice assistants like Alexa or Siri. They can also make suggestions in specific situations without the intervention of a human being and without being programmed for any such specific tasks. They are capable of unsupervised (and self-supervised) learning and maturing over a period of time thus providing better and more accurate results.

AI is first said to be in vogue in the mid-1950s, followed by Machine Learning (ML-machines that can learn from past incidents and old data) in 1980s. Thereafter, there was an *AI Winter* period due to lack of finance and interest in industry. However, with the availability of better processors and big-data in early 2000s things changed and *Deep Learning* (machines that can mimic a human brain) came about in 2010 followed by Generative AI (Gen-AI) in 2020, which is the latest development in the field of AI. Gen-AI means AI that can generate original intelligent content—such as long text (like as essay), realistic images, audio or video, etc, in response to a consumer demand, request or trigger (for example ChatGPT). Technologies that enable Gen-AI are ML and Deep Learning.[1,2] An important development in the field of AI is the *AI Agent* that is an autonomous system designed to sense the environment and take actions to attain defined objectives. These agents use Large Language Models (LLMs) for planning, execution, decision making and interaction with external environment; based on

environmental feedback the agent modifies the next iteration. Thus, in a cyclic manner, the agent is able to achieve the desired objective.

AI presents various advantages across businesses, industries and a variety of applications. Few important benefits are mechanisation of repetitive jobs, better analysis of huge amount of data in short span of time, improved decisions, lesser human-induced faults and reduced risks. Certain important use-cases of AI already being implemented are customer support, fraud sensing, personalised advertisements, application development and preventive maintenance of machinery. Important applications of AI include analysis of patients' data & medical imaging to diagnose diseases, fraud detection in finance, personalised customer experiences, inventory management and chat-bots. It is estimated that by 2030 AI will add about $15.7 trillion to the world economy.[3]

Like any new technology, AI also has its pitfalls like vulnerability to data poisoning or data corruption, unauthorised targeting of AI models by threat actors, AI model biases and breakdowns leading to system failures and vulnerabilities, gender/race biased decisions, privacy concerns, etc. It is well known that criminals generally adopt any new technology faster than anyone else. Same is true for AI also. A wide range of cybercrimes and cyber threats today involve extensive use of AI. On the other hand, AI can be an excellent tool for prevention of cybercrime and enhancing the cyber security. In this paper will discuss various means of exploiting AI for cybercrime as also the use of AI in fight against cybercrime.

Use of AI in Cybercrime

AI has altered the world of cybercrime and drastically changed the scope and scale of cybercrime. Using a tool like the ChatGPT for Gen-AI, a non-technical criminal can execute thousands of scams in a day sitting anywhere across the globe. AI is already in use for committing cybercrimes such as password cracking, voice cloning, deep fake videos and CAPTCHA breaks.[4] A UN study showed that the rise in AI based cybercrimes from 2023 to 2024 was about 1500% and Southeast Asia is fast emerging as a hub for AI based cybercrimes with 50% of deep-fake cybercrimes in Asia originating from Vietnam & Japan.[5] Common use cases of AI in execution of cybercrime are discussed further:

- **Deep Fake Video and Cloned Voice:** Using AI, fake videos, images, voice can be produced that are so close to the real thing that it is very convincing. Deep fakes are used to create explicit images/videos for blackmailing, bypass security features at secured premises like bank, identity theft, fraud, bypassing KYC requirements, spreading misinformation, tarnish reputation, etc. To achieve this, a Generative Adversarial Network (GAN) is used. GAN comprises of two competing neural networks called the *Generator* and the *Discriminator*. While the generator produces new data by using the data samples (of original voice, images and video), the discriminator analyses whether the newly created data is real or fake and gives this feedback to the generator. This goes on till the discriminator can no longer differentiate between the real and fake data. This way, using multiple iterations, GAN is able to generate fake data that is very close to the real data. Examples of GAN usage include music creation, image creation, enhancing image resolution, video creation, etc. Cloned voice can be used for social engineering attacks, mimicking a trusted person's voice. AI based voice manipulators can be used to anonymise telephone calls. Technologies like the *Dupleix16* of Google are capable of making phone calls with a human type voice instead of robot type voice. Open platforms like the Google's *WaveNet* can create video that imitates actual voice and facial expressions to such an extent that it is not discernible from the real one. In a study by the University of Waterloo, 39% of participants could not differentiate between real and fake videos.[6] There are several examples of cybercrimes committed using cloned voice and video in the last few years. Famous ones include a fake video of CFO of British firm Arup who participated in a video call with company employees and got them to transfer $25 million,[7] fake video of US president Joe Biden asking democrats in New Hampshire not to vote,[8] fake video of Elon Musk promoting a fake crypto exchange called *BitVex* with promise of very high returns, fake video of Ukrainian President surrendering to Russia, and fake video of Donald Trump asking for election donations in cryptocurrency.[9,10,11] A Delloitte study of 2024 found that 25.9% of companies did face a fake video attack and losses due to deep-fake could touch $40 billion by 2027.[12]

- **Social Engineering:** AI has considerably increased email attacks. AI tools can be used to write phishing emails that seem like genuine mails by taking care of

aspects such as wrong grammar, spelling mistakes, etc. Since the whole process is automated, these mails can be created and sent in bulk, with multiple variants, thereby improving the criminal's efficiency by automating large scale offensive cyber-attacks against individuals and/or organisations. AI based chatbots like *WormGPT* and *FraudGPT* are generally used for this purpose.

- **Password Cracking:** AI is being used to break passwords. Using AI based tools like the *PassGAN* and *Hashcat* cyber criminals can crack passwords with about 27% success rate.[13] By automating this process and carrying out millions of password combinations every second, the probability of success becomes higher.

- **No-Code AI Tools:** These are tools that can be used to develop applications without the knowledge of software coding. Individuals with no coding experience can utilise these tools to build malware codes or AI applications that can impersonate a person's writing style or build chat-bots that can behave like human beings. This brings in the concept of *AI Commoditisation*. Less trained cyber criminals can opt to utilise *AI-as-a-Service,* available to anyone willing to pay for the service.

- **Adversarial Attacks:** AI can manipulate the datasets to cause AI systems to give wrong, biased or skewed predictions and results. These are called *adversarial attacks* where data is intentionally manipulated to produce misleading results.

- **Intelligent Polymorphic Malware:** AI can be used to write self-modifying (polymorphic) malware which can detect the anti-malware software/tools on the target device and modify themselves to defy the tool and avoid anomaly detection by the device.

- **AI Assisted Ransomware Attacks:** Cyber criminals who resort to ransomware attacks by encrypting data can use AI to identify and predict vulnerabilities in a particular data hosting system and then exploit those gaps.

- **AI Assisted Vulnerability Search:** AI tools can be used to automatically search and find vulnerabilities in enterprise networks or data centres, and suggest possible attack vectors to exploit the weakness.

Use of AI in Cyber Defence

Like any new technology, AI also has its advantages that can be leveraged to fight cybercrime as well as make organisational cyber security more robust. AI, NLP and ML can be used in automatic detection, analysis and management of incident response, proactively detect and mitigate AI enabled cyber threats, real time analysis of network traffic, forecast of threats based on current and past traffic/incidents, etc. AI enables improved threat discovery, system robustness, adaptation, better identity management, controlled access mechanisms and automation of tasks. Various use cases of AI for such applications are discussed further:

- **Threat Intelligence and Predictive Mitigation:** AI can process a large amount of data in real time, simultaneously detecting patterns & anomalies and learning on the fly. AI can track data at scales beyond the capability of human beings in much lesser time. AI can analyse input data from multiple sources like network traffic, social media, dark web and intelligence threat feeds, etc simultaneously, and create advance awareness about impending threats. It effectively means that AI can detect Zero-Day threats. AI can even work out and suggest strategies to mitigate these future threats, thus preempting the threat actors. It also helps in reducing errors due to human fatigue, thus obviating false positives, false negatives and poor decisions. Anomalies can be detected automatically using ML, deep learning and NLP where large data sets including text, emails, chats, etc can be analysed in quick time frame. Behavioral analysis can be used to detect insider threats and Advanced Persistent Threats (APTs).[14,15,16]

- **Offensive Security:** AI enabled tools can simulate cyber-attacks on networks and systems to identify security vulnerabilities and report them to cyber security managers before attackers can exploit them. The whole process involves data collection from network traffic and open source, data collation & analysis, threat analysis and assessment, anomaly detection, attack simulation and finally identifying the vulnerabilities and reporting. This is like an automated Vulnerability Assessment & Penetration Testing (VAPT) analysis, which is presently being done manually.

- **Automated Incident Response:** In the case of a cyber incident taking place, AI enabled tools can automate a certain part of the response mechanism like isolation of affected systems from the network to contain the damage, alerting network administrators, etc.

- **Behaviour Detection:** AI can be helpful (using ML) to anticipate and respond to cyber threats by analysing behavioural patterns of cyber criminals.

- **Distributed Denial of Service (DDoS) Attack Detection:** AI powered neural networks can assimilate scattered information across the network and fuse it to ascertain if a DDoS attack is building up.

- **Bot Detection:** AI can be used to analyse web traffic and differentiate between good and bad bots. It can further proactively alert the cyber security team of such an attack enabling them to take timely actions to plug the vulnerabilities and deploy counter measures.

- **Malware Detection:** AI can be used to identify malware and viruses and over a period of time, the system can become more efficient and fast due to past experience and data.

- **AI Enabled Forensics:** AI forensic tools can analyse evidence, logs and past data to provide intelligent results that can help in post incident review and criminal proceedings. AI can also be used to assist criminal justice system by detection of criminals using real time face detection and image enhancing tools that can help to bring clarity to blurred images in CCTV footage.

- **Deep Fake and Cloned Voice Detection:** While AI is being used to create deep fake videos and cloned audio, it can also be used to detect them. AI identifies anomalies in these fake video/audio files by comparing with original samples and categorise the files as fake or real. It is like a cat and mouse game-same technology is being used to commit as well as detect crime.

- **Social Engineering and Spam Detection:** Phishing emails are the initial steps in building up an attack. AI can be used to scan emails for detecting phishing emails and spam by analysing the content, context, spelling mistakes, grammar mistakes, etc.

Challenges in AI Adoption

AI is a fairly new technology. Not many people understand it fully. The humankind is exploring a new world of possibilities with AI. AI is an emerging transformative force. Successful implementation of AI is contingent on many factors like good data-sets, adequate processing power, validating the results for accuracy, maturity of systems and so on. Apropos, AI implementation and adoption is riddled with challenges across technology, social, legal and ethical domains to include problems like data privacy and security, data accuracy, biases based on religion/gender/race, etc. Some of these important challenges are discussed below:

- **Finding the Correct Data Set:** As we know, AI systems work on huge data sets to train themselves and to produce outcomes. The AI driven output is therefore dependent on quality of this data set. Incorrect data sets may lead to not only wrong decisions but also biased decisions based on religion, race, gender, etc. This not only impacts the social fabric but is also an ethical and legal issue.

- **Data Security:** Data sets being huge in any AI system need commensurate data protection measures like encryption, safe storage, back-up, archival and retrieval. These are major challenges as they entail upgrading the in-house infrastructure. Captive infrastructure upgradation may not be cost effective with low returns of investment especially for small enterprise. Hence, businesses may resort to use of cloud infrastructure where the security implementation may not be meeting the desired standards. If data security is not implemented properly, it may also lead to violation of data privacy, which may give rise to ethical and legal concerns.

- **Requirement of High Compute Power:** AI requires substantially higher compute power compared to conventional systems. High performance processors like Graphical Processing Units (GPUs) and Tensor Processing Units (TPUs) are required instead of regular processors; these need more power also that adds to energy expenditure. These requirements can be challenging especially for small enterprises.

- **Limited Expertise:** Limited knowledge of engineers about AI systems is a very critical issue. It may give rise to irresponsible promotion and use of AI. The

situation gets more complex as it is very often difficult to understand the logic of AI systems reaching a particular decision. This kind of unexplainable nature of AI may also lead to lack of trust in AI systems.

- **Legal Challenges:** Several legal challenges arise out of AI processing and decision making. If AI produces an essay there will be issues related to copyright and intellectual property. If AI based decisions harm someone, accountability issues arise. If data is compromised, privacy issues will be flagged.

- **Ethical Issues:** Ethics in AI includes privacy issues, gender/race/religion-based discrimination, etc. These issues become more important in criminal justice system and healthcare where ethical issues take a front seat. Ethical issues also arise out of biases; if the data set fed to an AI system is biased, the system will learn from it and inherit such a bias. Such biases can lead to discrimination and unfair behaviour.

Conclusion

AI aims to make life easy and comfortable by developing systems and applications that can think and act like human beings. But, the misuse of same by criminals is on the rise. It is like a double-edged sword. In future, with AI becoming more mature and scaled-up, one could expect higher level of sophisticated cyber-attacks and cybercrimes which would be more difficult to detect and contain. Over reliance on AI based systems and tools may pose significant risks. Adoption of AI has also opened a barrage of challenges including ethical and legal issues. Existing legislations do not suffice to manage AI related aspects of any crime. There is also a need to have a legal definition of AI. While some like the European Union (EU) have already taken the lead in this regard, there is no specific regulation or legal framework to address the AI related issues in India. Aspects of liability are not very clear; for example, an AI application developed for a good cause may be misused by cyber criminals to commit a crime. In such case, probably the developer is also to be made accountable. Apropos, there is an urgent need to develop a legislation to deal with all these issues. Only a strong regulatory framework can make AI a good enabler for putting it to good use and deter its criminal exploitation.

DISCLAIMER

The paper is author's individual scholastic articulation and does not necessarily reflect the views of CENJOWS. The author certifies that the article is original in content, unpublished and it has not been submitted for publication/ web upload elsewhere and that the facts and figures quoted are duly referenced, as needed and are believed to be correct.

Endnotes

[1] Stryker, Cole, 09 Aug 2024, *What is artificial intelligence*, available at https://www.ibm.com/think/topics/artificial-intelligence, accessed 29 Jan 2025

[2] IIT Kanpur, 14 Oct 2024, *What is Artificial Intelligence in Simple Words*, available at https://eicta.iitk.ac.in/knowledge-hub/artificial-intelligence/what-is-artificial-intelligence/, accessed 29 Jan 2025

[3] *PwC's Global Artificial Intelligence Study: Exploiting the AI Revolution*, 2021, available at https://www.pwc.com/gx/en/issues/artificial-intelligence/publications/artificial-intelligence-study.html, accessed 04 Feb 2025

[4] Scott Monteith, Tasha Glenn, John R. Geddes, Eric D. Achtyes, Peter C. Whybrow and Michael Bauer, 23 Sep 2024, *Artificial Intelligence and Cybercrime: Implications for Individuals and the Healthcare Sector*, The British Journal of Psychiatry, Vol 225, Issue 4, available at https://www.cambridge.org/core/journals/the-british-journal-of-psychiatry/article/artificial-intelligence-and-cybercrime-implications-for-individuals-and-the-healthcare-sector/6409A9AB77FE31DD8033D7B761D20381, accessed 29 Jan 2025

[5] https://www.darkreading.com/threat-intelligence/ai-powered-cybercrime-cartels-asia, accessed 19 Feb 2025

[6] https://incode.com/blog/top-5-cases-of-ai-deepfake-fraud-from-2024-exposed/, accessed 19 Feb 2025

[7] Ibid

[8] Ibid

[9] Jabeen, Ansari Zartab, 2024, *Camouflage of AI in Cyber Crimes Vis-a Vis legal issues and Challenges*, available at https://woxsen.edu.in/woxsen-law-review/wlr-papers/camouflage-of-AI-in-cyber-crimes-vis-a-vis-legal-issues-and-challenges/#:~:text=Role%20of%20AI%20in%20committing,and%20AI%20is%20no%20different%E2%80%9D, accessed 29 Jan 2025

[10] Team Blink, 26 Aug 2024, *5 Ways Cybercriminals Are Using AI in Cybercrime in 2024*, available at https://www.blinkops.com/blog/using-ai-in-cybercrime, accessed 29 Jan 2025

[11] Staveley Confidence , 24 Aug 2024, *AI in Cybersecurity – Q2 2024 Insights*, available at https://aicyberinsights.com/ai-in-cybersecurity-q2-2024-insights/, accessed 03 Feb 2025

[12] N 6

[13] N 9

[14] ThreatMon, 2024, *AI-Powered Threat Intelligence: A Comprehensive Handbook*, available at https://threatmon.io/ai-powered-threat-intelligence-a-comprehensive-handbook/, accessed 03 Feb 2025

[15] Kumar Atul and Raman Anand, DSCI Report, 2021, *India's AI/ML Cybersecurity Capabilities*, available at https://www.dsci.in/resource/content/indias-ai-ml-cybersecurity-capabilities-driving-innovation-cybersecurity-front, accessed 15 Nov 2024

[16] DSCI Report, 2021, *India's AI/ML Cybersecurity Capabilities*, available at https://www.dsci.in/resource/content/indias-ai-ml-cybersecurity-capabilities-driving-innovation-cybersecurity-front, accessed 15 Nov 2024