BJPsych | The British Journal of Psychiatry (2024)
Page 1 of 3. doi: 10.1192/bjp.2024.77

BJPsych Editorial

Artificial intelligence and cybercrime: implications for individuals and the healthcare sector

Scott Monteith, Tasha Glenn, John R. Geddes, Eric D. Achtyes, Peter C. Whybrow and Michael Bauer

Summary

The malicious use of artificial intelligence is growing rapidly, creating major security threats for individuals and the healthcare sector. Individuals with mental illness may be especially vulnerable. Healthcare provider data are a prime target for cybercriminals. There is a need to improve cybersecurity to detect and prevent cyberattacks against individuals and the healthcare sector, including the use of artificial intelligence predictive tools.

Keywords

Cybercrime; artificial intelligence; health informatics; mental health services; risk assessment.

The malicious use of artificial intelligence has created new types of security threat for both individuals and the healthcare sector. Although artificial intelligence is a fundamental technology of our age, it has enabled the creation of new types of large-scale cyberthreat, and artificial intelligence-based cybercrime has grown rapidly worldwide. Medical data are a prime target for cybercriminals, given the high value of stolen data. Of further concern to psychiatry is that both patients with mental illness and mental health data may be especially vulnerable to artificial intelligence-based security threats. This editorial will discuss the common types of artificial intelligence-based cyberthreat faced by both individuals and the healthcare sector and address potential ways to mitigate risks, including the use of artificial intelligence predictive tools.

Artificial intelligence as a security threat

Artificial intelligence has transformed cybercrime. Cybercriminals are using artificial intelligence to enhance attacks so that it is harder for antivirus software to detect, to create new types of attack based on synthetic data (deepfakes) and to automate the creation of large-scale attacks. Artificial intelligence has changed the scope of fraudulent schemes. Today, a single scammer can use generative artificial intelligence such as ChatGPT to run hundreds of thousands of scams 24 h a day from anywhere in the world.[1] Generative artificial intelligence has made it easier for criminals with only a limited programming skills or technical knowledge. The phishing text provided by large language models (LLMs) is more sophisticated and without spelling and grammatical errors, unlike spam emails of the past.[1] Criminals can also purchase data from data brokers to help customise phishing attacks. Additionally, there are malicious LLMs, WormGPT and FraudGPT, developed specifically for criminals and advertised on underground forums. Generative artificial intelligence is also used by hackers to crack passwords, with most being cracked in less than a minute.

Artificial intelligence as a security threat for individuals

Individuals face many types of artificial intelligence-enabled cyberattack. Cybercriminals use artificial intelligence voice-cloning technology to impersonate people and convince family members to send money.[2] Artificial intelligence-based facial recognition has magnified the ability to recognise individuals. Artificial intelligence is used to manipulate photos and videos to create explicit content for extortion schemes. Many artificial intelligence-based phishing attacks lead victims to a site to harvest passwords and other personal credentials. Artificial intelligence can be used to replicate writing styles and impersonate users such that fake messages are difficult to distinguish from genuine communications. Artificial intelligence applications can track user activities to learn habits and preferences. Cybercriminals are enticing people to invest in fraudulent schemes by claiming that artificial intelligence is involved.

Artificial intelligence as a security threat for individuals with mental illness

The use of artificial intelligence to scam individuals is of particular concern for psychiatry, as mental illness may increase the vulnerability to cybercrime. Factors that may increase vulnerability to online deception include severe mental illness, emotional instability, poorer short-term memory, cognitive impairment and a lack of technical skills. Additionally, impulsivity and overconfidence in information technology knowledge and skills may increase susceptibility to cybercrime. Psychiatrists should be aware that artificial intelligence is increasing the quantity and quality of cybercrime against individuals and be able to recommend trusted sources for consumer education on cybersecurity to their patients. Individuals of all ages, backgrounds and levels of technological sophistication need to learn the best practices to protect themselves from cyberattacks.

Artificial intelligence as a security threat for healthcare providers

Businesses also face many types of artificial intelligence-enabled cyberattack. Artificial intelligence-based attacks against business to steal money and critical information include impersonation and targeted phishing attacks against employees, more effective ransomware, distributed denial of service attacks and attacks that hijack cloud infrastructure.[3] In healthcare, artificial intelligence is increasingly used for a variety of purposes, including customer service, administrative tasks, diagnosis in radiology and pathology, ongoing patient monitoring, drug discovery and medical research. Although artificial intelligence systems are increasingly used for critical applications, some in management may not realise that artificial intelligence systems are vulnerable to cyberattacks. Adversarial attacks, which involve input data

https://doi.org/10.1192/bjp.2024.77 Published online by Cambridge University Press

intentionally crafted to cause misclassification by an artificial intelligence model, are of major concern across many domains, including medicine.[4] Incorrect classification from an adversarial attack may cause false diagnostic predictions, as demonstrated with medical imaging.

An adversarial attack may degrade the performance of a healthcare system collecting data from multiple connected smart medical devices and may target susceptible locations in electronic medical records (EMRs). In other studies, the targets of adversarial attacks include an electrocardiogram, an implantable cardioverter defibrillator and an electroencephalogram. Adversarial attacks may also occur against speech-based emotion recognition systems. Medical image models may be more vulnerable to adversarial attacks than natural image models owing to specific characteristics of medical image data and models.

Need to monitor artificial intelligence security risks

There is a critical need for robust security to monitor personal and business data against artificial intelligence cybercrime. Consumers need increased awareness of cybercrime and ongoing training on how to protect against it. Mental illness may increase vulnerability to online fraud, and victimisation may worsen the symptoms of mental illness. Some characteristics associated with increased vulnerability to online fraud include impulsivity, older age, sensation-seeking, cognitive impairment, willingness to trust and a lack of technical knowledge. The effects of cybercrime on victims are long-lasting, with impacts that are psychological as well as financial.

Artificial intelligence used to monitor security risks

Business also faces new challenges as security approaches must be able to detect artificial intelligence attacks that an organisation often has never seen before. The increasing use of artificial intelligence for critical applications in businesses, including healthcare, creates greater incentives for cybercriminals to attack these algorithms and increases the negative consequences of a successful attack. On an enterprise level, businesses, including healthcare, need to take a robust, multifaceted approach that includes artificial intelligence-driven cybersecurity solutions to enhance the more traditional human and technology approaches to combat cybercrime. Artificial intelligence can provide continuous monitoring, recognise and diagnose threats in real time, help identify false positives and improve access control management. Artificial intelligence can detect pattern changes in the overall data ecosystem with a level of sensitivity that would not be recognised by humans. Artificial intelligence should be involved in the many technological challenges of providing cybersecurity in the increasingly complex and interconnected environments. Artificial intelligence cybersecurity tools may detect attacks to remote patient monitoring devices which now routinely involve artificial intelligence.[5] Artificial intelligence tools are needed to provide a comprehensive approach to defend against artificial intelligence cyberattacks. The use of artificial intelligence cybersecurity tools is of particular importance in healthcare, given the potential for system failures to cause severe harm to individuals.

Limitations of this editorial

The limitations of using artificial intelligence for cybersecurity, including details of the methods involved in artificial intelligence-based detection, and prediction of threats and malicious activities, are not discussed here. The financial investment required for artificial intelligence-based cybersecurity, and the costs of acquiring huge volumes of training data, are not estimated. Other types of cybercrime are not discussed, including risks to artificial intelligence algorithms. Policy proposals for the governance of artificial intelligence algorithms, issues of transparency, auditing and accountability are not suggested.

Conclusion

Artificial intelligence-based security threats are a serious concern for both individuals and the healthcare sector. Individuals with mental illness may have an increased susceptibility to artificial intelligence enabled cyberattacks. Healthcare records, including mental health, are a prime target. Numerous incidents, including adversarial attacks, have occurred against individuals and the healthcare system. As both dependence on technology and the sophistication of cybercriminals has increased, there is a need to augment cybersecurity with artificial intelligence to detect and prevent cyberattacks. Artificial intelligence-based cybersecurity is an important and necessary addition to cybersecurity across domains, including healthcare.

Scott Monteith, Department of Psychiatry, Michigan State University College of Human Medicine, Traverse City Campus, Traverse City, Michigan, USA; **Tasha Glenn**, ChronoRecord Association, Fullerton, California, USA; **John R. Geddes**, Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford, UK; **Eric D. Achtyes**, Department of Psychiatry, Western Michigan University Homer Stryker M.D. School of Medicine, Kalamazoo, Michigan, USA; **Peter C. Whybrow**, Department of Psychiatry and Biobehavioral Sciences, Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles (UCLA), Los Angeles, California, USA; **Michael Bauer**, Department of Psychiatry and Psychotherapy, University Hospital Carl Gustav Carus Faculty of Medicine, Technische Universität Dresden, Dresden, Germany

Correspondence: Scott Monteith. Email: monteit2@msu.edu

First received 17 Nov 2023, accepted 25 Mar 2024

Supplementary material

Supplementary material is available online at https://doi.org/10.1192/bjp.2024.77.

Data availability

Data availability is not applicable to this article as no new data were created or analysed in this study.

Author contributions

S.M. and T.G. wrote the initial draft. All authors reviewed and approved the final manuscript.

Funding

This work received no specific grant from any funding agency, commercial or not-for-profit sectors.

Declaration of interest

J.R.G. is a member of the *BJPsych* editorial board and did not take part in the review or decision-making process of this paper.

References

1 Schneier B, Raghavan B. LLMs and phishing. *Schneier on Security Blog*, 2023; 10 Apr (https://www.schneier.com/blog/archives/2023/04/llms-and-phishing.html).

2 McAfee. *Beware the Artificial Impostor: A McAfee Cybersecurity Artificial Intelligence Report*. McAfee, 2023 (https://www.mcafee.com/en-us/resources/cybersecurity-reports-and-guides.html).

https://doi.org/10.1192/bjp.2024.77 Published online by Cambridge University Press

3 MIT Technology Review Insights. *Preparing for artificial intelligence-Enabled Cyberattacks*. MIT Technology Review Insights, 2021 (https://www.technologyreview.com/2021/04/08/1021696/preparing-for-ai-enabled-cyberattacks/).

4 Finlayson SG, Bowers JD, Ito J, Zittrain JL, Beam AL, Kohane IS. Adversarial attacks on medical machine learning. *Science* 2019; **363**: 1287–9.

5 Vijayakumar KP, Pradeep K, Balasundaram A, Prusty MR. Enhanced cyber attack detection process for Internet of Health Things (IOHT) devices using deep neural network. *Processes* 2023; **11**: 1072.

Additional references are included in the Supplementary material available at https://doi.org/10.1192/bjp.2024.77.

EXTRA CONTENT ONLINE

https://doi.org/10.1192/bjp.2024.77 Published online by Cambridge University Press