UDC 004.8:004.056

AI IN CYBERSECURITY

Danilevich P.V.

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Malikova I.G. – Senior Lecturer at the Department of Foreign Languages

Annotation. Artificial Intelligence (AI) is rapidly revolutionizing the field of cybersecurity serving as both a powerful tool of protection against cybercrimes and new possibilities for criminals and scammers. This article is devoted to the impact of artificial intelligence on cybersecurity, examining both defensive innovations and emerging AI-powered threats.

Keywords: Artificial Intelligence (AI), cybersecurity, cyber attacks, machine learning, vulnerabilities.

Introduction. At present, technologies are developing very quickly, and so are cyber threats. Cyber criminals now use sophisticated methods and modern technologies to attack systems, steal or damage data, making traditional security measures less effective. Artificial intelligence offers new solutions to this problem by enhancing the ability to detect and respond to cyber threats. However, the same AI technologies can also be exploited by criminals.

Main part. Before AI, creating a convincing phishing email required good writing skills, cultural knowledge and manual effort. Now, anyone can use AI tools to write perfect fake messages in any language. This has led to a massive increase in attacks. The scale of this transformation is overwhelming. Phishing attacks have increased exponentially following the widespread availability of generative AI tools, with a significant majority of phishing emails now incorporating AI in some form. Figure 1 provides a comparative overview of the key differences between traditional human-only phishing and modern AI-assisted phishing campaigns [1].

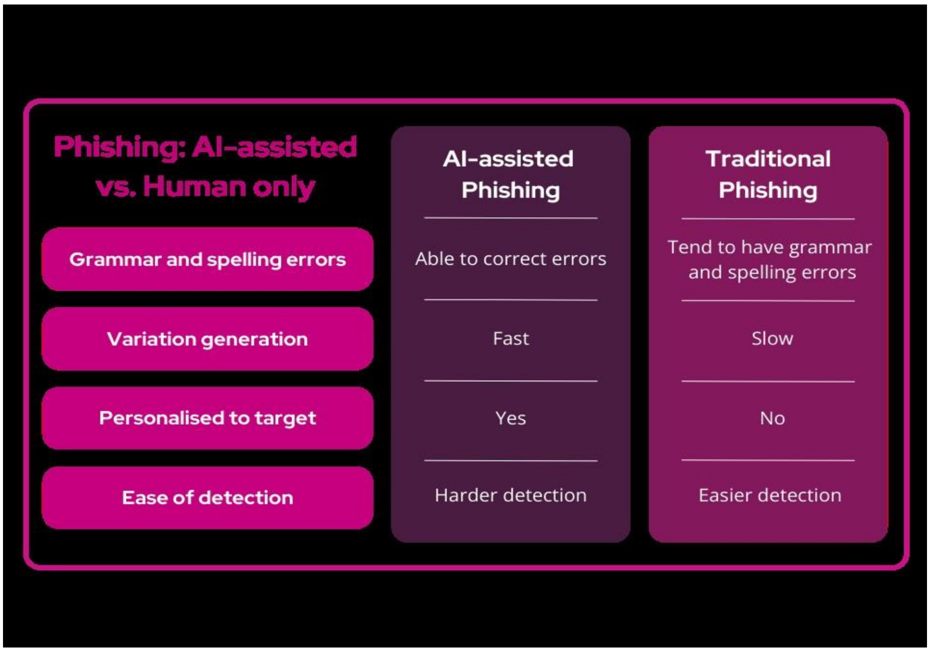

Figure 1 – Comparative Analysis of Traditional vs. AI-Assisted Phishing Attacks

Deepfake technology has introduced entirely new threat vectors that fundamentally challenge traditional authentication methods. Cybercriminals now use Artificial intelligence to create convincing voice replicas and deepfake videos of company leaders. By impersonating executives, they manipulate unsuspecting employees into approving unauthorized wire transfers, disclosing confidential information, or bypassing standard security protocols, resulting in significant financial losses. The FBI’s Internet Crime Complaint Center reported a 37% rise in AI-assisted business email compromise schemes during 2025, with hundreds of documented deepfake-based scams [2].

Attackers are increasingly using AI to discover vulnerabilities in target systems. Machine learning models can scan codebases, analyze network configurations, and identify weaknesses at scales impossible for human researchers. These AI-powered tools can autonomously probe for zero-day exploits and prioritize the most prospective attack vectors.

The rapid adoption of generative AI tools in enterprise environments has created new data leak risks. Employees pursuing productivity gains often adopt AI tools without organizational approval or oversight. This phenomenon is called “shadow AI”. When employees input private data, source code, or sensitive documents into public AI models, that information may be exposed to unauthorized persons. The consequences of shadow AI extend beyond immediate data leak. Once confidential information enters public AI models, organizations lose control over that data permanently. It may appear in responses to other users, be incorporated into model updates, or be extracted through adversarial prompting techniques [3].

However, Artificial intelligence is used not only for fraudulent or other illegal purposes. AI helps security teams detect attacks faster. Unlike old systems that could only find known threats, AI can recognize unusual activity even if it’s never been seen before. It learns over time and gets better at telling the difference between real threats and false alarms.

AI is good at analyzing emails to spot phishing attempts. Modern AI systems analyze message tone, content, structure, and metadata to identify sophisticated phishing that bypasses conventional filters.

Sometimes the danger comes from inside the company either from people whose accounts have been hacked. AI learns what “normal” behavior looks like for each user. If someone suddenly starts accessing files at 3 a.m. or logging in from a strange location, AI flags it for investigation.

Companies that use AI for security save money. On average, they save about $1.9 million compared to companies that don’t use AI. One study found that AI detects 98% of threats and reduces response time by 70%. Another study showed AI achieving 99.95% accuracy in detecting intrusions. By minimizing manual effort, AI has helped speed up numerous processes in data security. It enables security teams to quickly identify and classify sensitive information across the entire environment, whether it resides on the company’s local infrastructure or in a cloud application. AI can also instantly detect attempts to move data outside the organization, either blocking the action in real time or flagging it for the security team [4].

Conclusion. The integration of Artificial intelligence into cybersecurity represents one of the most significant transformations in the history of information technology. AI is a double-edged sword in cybersecurity. AI-powered defenses offer improvements in detection accuracy and response speed. Machine learning and deep learning systems can identify novel threats and analyze code for vulnerabilities.

Simultaneously, AI is changing the threat landscape. The same AI tools that help protect companies are also helping attackers. Cybercriminals are now using AI to automate their attacks, make them more convincing, and scale them up in ways that were never possible before. This is the dual-use problem of AI-the fact that the same technology can be used for both good and bad purposes.

This dual-use nature of AI creates an arms race between defenders and criminals. Artificial intelligence is not just another tool in the arsenal, but a powerful force that is fundamentally changing the field of cybersecurity. The main challenge for the near future is to learn how to use it as effectively as possible for defense, while preventing it from turning into a dangerous weapon.

References

1. Sekuro.io. Emerging Threat: AI vs. Human Deceit [Electronic resource]. Mode of access: https://sekuro.io/blog/emerging-threat-ai-vs-human-deceit/. Date of access: 01.03.2026.

2. DeepStrike. AI Cyber Attack Statistics 2025, Trends, Costs, Defense [Electronic resource]. Mode of access: https://deepstrike.io/blog/ai-cyber-attack-statistics-2025/. Date of access: 20.02.2026.

3. Nesktope. Cloud and Threat Report: 2026 [Electronic resource]. Mode of access: https://www.netskope.com/resources/cloud-and-threat-reports/cloud-and-threat-report-2026. Date of access: 03.03.2026.

4. Microsoft. What is AI for cybersecurity? [Electronic resource]. Mode of access: https://www.microsoft.com/en-us/security/business/security-101/what-is-ai-for-cybersecurity#History. Date of access: 20.02.2026.

5. AI and the Future of Cybersecurity [Electronic resource]. Mode of access: https://extension.harvard.edu/blog/ai-and-the-future-of-cybersecurity/#How-AI-Enables-the-Next-Generation-of-Cyber-Attacks-. Date of access: 21.02.2026.