Dominika Grzybowska-Ganszczyk, PhD
Jerzy Kukuczka Academy of Physical Education in Katowice
e-mail: dominikagrzybowska@yahoo.com
ORCID: 0000-0002-6413-306X

DOI: 10.26410/SF_1/26/6

ARTIFICIAL INTELLIGENCE AS A THREAT MULTIPLIER IN CRIMINAL ACTIVITY AND INFLUENCE OPERATIONS. THE ALGORITHMIC ARMS RACE IN INTERNAL SECURITY

Abstract

Artificial intelligence (AI) is becoming one of the key factors transforming the contemporary internal security environment of states. Its dynamic development not only enhances the effectiveness of analytical and operational tools used by public institutions, but also significantly lowers the entry barrier for individuals engaged in criminal activities and for actors seeking to influence public opinion and social behaviour. In particular, there is a growing use of AI technologies in digital crime and destabilizing information activities, including disinformation, message manipulation, and advanced forms of social engineering. These phenomena contribute to a qualitative shift in the nature of threats—from individual, manual actions to automated, scalable systems capable of generating content and interacting with users. The aim of this article is to analyse artificial intelligence as a "threat multiplier" in four key areas: the generation of synthetic multimedia materials (deepfakes), the automation of phishing attacks, the mass production of disinformation, and precise, personalized social engineering attacks.

At the same time, defensive AIbased systems are being developed in parallel by state institutions and international organizations to detect, analyse, and neutralize information threats. This dynamic leads to what is increasingly described as an "algorithmic arms race," in which strategic advantage depends to a growing extent on dataprocessing capabilities, model quality, and the speed at which systems can adapt.

Keywords

artificial intelligence, internal security, digital crime, algorithmic arms race, cybersecurity

Introduction

Contemporary security studies increasingly describe the information environment as a fully fledged and autonomous domain of competition and conflict, alongside the traditional land, maritime, air, and cyber domains. Within this framework, the flow of information, its credibility, and the manner in which it is processed become significant factors determining the stability of social and institutional systems. Against this background, artificial intelligence (AI) ceases to function solely as a tool supporting technological processes. Increasingly, AI acts as a catalyst for changes in the structure of contemporary threats, influencing both their scale and nature. Institutions responsible for cybersecurity – such as European Union Agency for Cybersecurity – as well as law enforcement agencies, including Europol, emphasize that the development of AI systems significantly lowers the entry threshold for criminal activities while simultaneously increasing their automation, scalability, and resistance to detection.[1]

The development of digital technologies in the twenty-first century has led to a profound transformation in the functioning of states, economies, and societies. Information systems, telecommunication networks, and AI-based solutions have become integral components of critical infrastructure and social life. At the same time, this process has generated new and complex areas of risk, the dynamics of which extend beyond traditional security analysis frameworks. Contemporary cybercrime increasingly assumes more complex, organized, and transnational forms. It exploits the anonymity of the online environment, encryption mechanisms, and the automation of attack processes, enabling perpetrators to conduct large-scale operations at relatively low operational costs. Alongside classical forms of digital crime, such as phishing[2] and ransomware, new-generation

1 ENISA. (2023). ENISA Threat Landscape 2023. European Union Agency for Cybersecurity.

2 Jason Hong. (2012). The State of Phishing Attacks. Communications of the ACM, 55(1), 74–81. https://doi.org/10.1145/2063176.2063197

phenomena have emerged, including deepfake[3], pharming[4], and sextortion[5], which directly affect information security, social trust, and the credibility of public communication.

A consequence of this evolution is the growing burden placed on institutions responsible for internal security. Law enforcement agencies and intelligence services increasingly encounter the limitations of traditional operational and investigative methods, which are not always adequate in relation to the dynamics of the digital environment. Particular challenges concern the identification of perpetrators operating in transnational cyberspace, the analysis of large datasets, and shortages of specialized competencies in the field of digital analysis.

As a result, cybercrime, once regarded as a marginal phenomenon, has become a permanent and structural element of the contemporary security environment. Its global nature renders traditional jurisdictional and operational models progressively obsolete, thereby necessitating the search for new analytical and operational tools.

In response to these challenges, AI-based systems are playing an increasingly important role in both criminal analysis and preventive activities. AI enables the processing of large datasets, the identification of patterns of criminal behavior, and support for predictive processes in the field of risk assessment. Simultaneously, its application raises significant legal and ethical questions, particularly regarding the limits of decision-making automation, the protection of privacy, and compliance with the principles of legality and proportionality in state activities. Accordingly, the principal research problem of this study focuses on the analysis of the possibilities for using artificial intelligence in combating cybercrime and its impact on the operational practices of law enforcement agencies. The article adopts an analytical and descriptive approach and encompasses technological, operational, and legal aspects.

The objective of this study is to demonstrate the potential of AI as a tool supporting internal security while simultaneously identifying the limitations and dilemmas arising from its implementation in institutional practice. Consequently, the need to develop a balanced approach is emphasized, based on maintaining equilibrium between the effectiveness of operational activities and the protection of individual rights as well as the standards of the rule of law.

3 Robert Chesney & Danielle K. Citron, Deep Fakes: a Looming Challenge for Privacy, Democracy, and National Security, 107 California Law Review 1753 (2019). Available at: https://scholarship.law.bu.edu/faculty_scholarship/640

4 Kolla, Jathin & Praneeth, Shinde & Baig, Mirza & Reddy, K.. (2022). a comparison study of machine learning techniques for phishing detection. Journal of Business and Information Systems (e-ISSN: 2685-2543). 4. 21-33. 10.36067/jbis.v4i1.120

5 Europol. (2017). Online sexual coercion and extortion as a form of crime affecting children: Law enforcement perspective (updated 2021). Europol

Artificial intelligence as a tool of digital crime and the transformation of the cyber threat environment

Contemporary cybercrime constitutes one of the most dynamic and adaptive phenomena within the global security environment. Its development remains closely correlated with technological progress, which not only provides new protection tools but simultaneously redefines the operational capabilities of criminal actors. Consequently, digital crime ceases to be a collection of isolated incidents and increasingly assumes the form of complex, automated, and transnational operational systems.[6]

The academic literature increasingly emphasizes that cybercrime has attained a systemic dimension, generating losses amounting to trillions of dollars annually and directly affecting the stability of economies and state security. Recent analyses indicate that the global costs of cyber incidents encompass not only breaches of data confidentiality, theft of financial resources, or loss of intellectual property, but also long-term operational disruptions and substantial costs associated with infrastructure recovery following attacks. Forecasts by Cybersecurity Ventures indicate that by 2025 the total value of damages caused by cybercrime may reach USD 10.5 trillion annually, representing an unprecedented increase compared with 2015 (Cybersecurity Ventures, 2020).[7]

Simultaneously, a dynamic professionalization of the criminal environment can be observed, in which the dominant operational model is becoming Crime-as--a-Service. Criminal groups increasingly function as specialized service providers, offering infrastructure, anonymization tools, exploit kits, botnets, or ready-made phishing campaigns. This phenomenon significantly lowers the entry threshold for new perpetrators, enabling them to conduct advanced operations without possessing specialized technical knowledge. Europol and European Union Agency for Cybersecurity have for several years indicated that the development of the service-based cybercrime model constitutes one of the key factors driving the increase in the number of incidents and their automation.

A particularly significant area of threat escalation is ransomware, which over the last decade has transformed from a tool of individual attacks into a global criminal ecosystem. a report by Cybersecurity Ventures (2023) forecasts that annual ransomware-related costs may rise to USD 265 billion by 2031. This increase results from the growing activity of state-sponsored groups, increasingly sophisticated extortion techniques, and the automation of the entire attack chain – from vulnerability scanning to negotiations with victims. In response to the scale of the threat, many states have classified ransomware as one of the most serious

6 Maria Bada, Jason R.C. Nurse, Developing cybersecurity education and awareness programmes for small- and medium-sized enterprises (SMEs), Information and Computer Security, Volume 27, Issue 3, 2019, Pages 393 410, ISSN 2056-4961, https://doi.org/10.1108/ICS-07-2018-0080.

7 Kshetri, Nir. (2021). Cybersecurity Management: An Organizational and Strategic Approach. 10.3138/9781487531249.

challenges to national security, requiring coordinated legislative, operational, and technological measures.[8]

The strategic significance of ransomware as a security threat is also reflected in the private sector. According to the Verizon Data Breach Investigations Report (2023), ransomware attacks and other forms of cybercrime have the greatest impact on public administration, the IT sector, finance, manufacturing, and professional services. This indicates that cyber threats are no longer a marginal phenomenon but rather an element of permanent operational risk, requiring the implementation of approaches based on risk analysis, operational resilience, and the continuous improvement of detection mechanisms.[9],[10]

In this context, artificial intelligence plays a particularly significant role, becoming a factor that qualitatively transforms the nature of threats. This transformation does not merely concern the enhancement of existing operational methods, but rather the emergence of entirely new possibilities for generating, modifying, and distributing content and interactions with criminal potential. AI thus enters the sphere of criminal activity as an infrastructural rather than solely instrumental element.[11]

One of the most visible manifestations of this process is the development of generative technologies enabling the creation of synthetic audio-visual materials, commonly referred to as deepfakes.[12] These systems allow for the realistic reproduction of the image and voice of public figures, including representatives of state institutions, opinion leaders, and public officials. From the perspective of internal security, the significance lies not only in the possibility of falsifying content, but above all in its systemic consequence, described as the erosion of cognitive trust.

This phenomenon consists in the gradual weakening of recipients' ability to distinguish reliable information from synthetically generated content. As a result, the foundations of social communication become destabilized, to the extent that even authentic materials may be questioned as potentially manipulated.[13] Consequently, deepfakes cease to function merely as tools of deception and instead become instruments capable of influencing social perception itself.[14]

---

8 Jimmy, Fnu. (2023). Understanding Ransomware Attacks: Trends and Prevention Strategies. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online). 2. 180-210. 10.60087/jklst.vol2.n1.p214.

9 Edwards, Benjamin & Hofmeyr, Steven & Forrest, Stephanie. (2015). Hype and Heavy Tails: a closer look at data breaches.

10 Romanosky, Sasha. (2016). Examining the costs and causes of cyber incidents. Journal of Cybersecurity. 2. tyw001. 10.1093/cybsec/tyw001.

11 Singh, S., & Dhumane, A. (2025). Unmasking Digital Deceptions: An Integrative Review of Deepfake Detection, Multimedia Forensics, and Cybersecurity Challenges. MethodsX, 15, 103632. https://doi.org/10.1016/j.mex.2025.103632

12 Whyte, C. (2020). Deepfake news: AI-enabled disinformation as a multi-level public policy challenge. Journal of Cyber Policy, 5(2), 199–217. https://doi.org/10.1080/23738871.2020.1805482

13 Gibson, P., & Connolly, S. (2023). Fit for purpose? Taking a closer look at the UK's Online Media Literacy Strategy. Journal of Media Literacy Education, 15(1), 109-115.

14 Shoaib, M. R., Wang, Z., Ahvanooey, M. T., & Zhao, J. (2023, November). Deepfakes, misinformation, and disinformation in the era of frontier AI, generative AI, and large AI models. In 2023 international conference on computer and applications (ICCA) (pp. 1-7). IEEE.

Simultaneously, artificial intelligence reinforces classical forms of cybercrime, such as phishing, ransomware, pharming, and sextortion. The key transformation does not concern the typology of threats itself, but rather their operational quality. The automation of content-generation processes enables the creation of multilingual, personalized, and dynamically adaptive communications, significantly increasing the effectiveness of social engineering attacks.[15] In practice, this signifies a transition from mass, non-selective campaigns to precise, targeted operations in which the analysis of victim-related data enables narratives to be tailored to behavioral and emotional characteristics. In this sense, AI functions as an amplifier of social engineering, increasing both the scale and effectiveness of influence.[16]

An important element of the contemporary cyber threat environment is also the disappearance of the physical crime scene. In many cases, perpetrators have no direct contact with victims, and the entire criminal process is conducted within the digital environment. Examples of so-called cyber kidnappings demonstrate that psychological and communicative manipulation may lead to real material consequences despite the absence of the perpetrator's physical presence. This mechanism illustrates the shift of criminal activity from the physical domain to the cognitive and informational sphere.[17]

Additionally, increasingly widespread use of botnets and automated attack systems can be observed, enabling simultaneous large-scale operations with minimal human involvement.[18] Within this model, artificial intelligence performs a coordinating function, optimizing target selection, attack timing, and methods of execution. Consequently, cybercrime is beginning to assume an infrastructural form in which individual elements – from content generation, through data analysis, to attack execution – are integrated within a single technological ecosystem.[19] Such a structure significantly complicates detection and neutralization efforts, as it is not based on isolated activities but rather on a continuous adaptive process.[20]

In recent years, particular significance has been attributed to synthetic forms of abuse, including the generation of false compromising materials, such as deepfake pornography. Such content constitutes not only a tool of blackmail but also an instrument of long-term reputational damage and erosion of social trust, the consequences

15 Schmitt, Marc & Flechais, Ivan. (2024). Digital deception: generative artificial intelligence in social engineering and phishing. Artificial Intelligence Review. 57. 10.1007/s10462-024-10973-2.

16 Modupe Awotidebe (2023). The Rise of Intelligent Threats: Exploring AI-Driven Cybercrime in the Digital Era.

17 Chang, L. Y.-C., Zhou, Y., & Phan, D. H. (2023). Virtual kidnappings: Online scams with "Asian characteristics" during the pandemic. In R. G. Smith, R. Sarre, L. Y.-C. Chang, & L. Y.-C. Lau (Eds.), Cybercrime in the pandemic, the digital age and beyond. Palgrave Macmillan. https://doi.org/10.1007/978-3-031-29107-4_6

18 Kolias, C., Kambourakis, G., Stavrou, A., & Voas, J. (2017). DDoS in the IoT: Mirai and other botnets. computer, 50(7), 80-84.

19 Mahboubi, A., Luong, K., Aboutorab, H., Bui, H. T., Camtepe, S., Ansari, K., & Barry, B. I. A. (2025). The evolving botnet threat landscape: a comprehensive analysis of detection techniques in the age of artificial intelligence. Internet of Things, 33, 101728. https://doi.org/10.1016/j.iot.2025.101728

20 Pum, Mengkorn. (2022). The Rise of AI-Powered Cybercrime: Implications for Digital Security and Policy.

of which may be enduring and difficult to reverse.[21] All these phenomena generate substantial implications for law enforcement agencies and internal security systems. Traditional investigative methods, based on the analysis of physical traces and linear procedural models, prove insufficient in an environment where data are dispersed, dynamic, and automatically generated. This necessitates the development of AI-based analytical tools capable of processing large datasets in near real time and identifying hidden behavioral patterns.[22] As a result, the contemporary cybercrime environment may be described as an adaptive system in which technology, automation, and psychological mechanisms form a coherent structure of threats. Within this framework, artificial intelligence is not merely a supporting tool but a factor that structurally transforms the functioning of digital crime.[23]

Implications for state internal security

The transformation of the threat environment associated with the development of artificial intelligence leads to a fundamental change in the functioning of internal security systems. This is no longer solely a matter of modernizing operational tools, but rather a necessity to redefine the fundamental assumptions underlying the perception of threats, evidence, and institutional responsibility.[24]

An important area of transformation within the contemporary digital security environment is the profound redefinition of the concept of evidence in the information sphere. Under conditions of widespread availability of generative technologies – including both advanced deepfake systems and synthetic voice reconstruction – the traditional understanding of evidentiary material as a faithful representation of reality is fundamentally challenged. Digital artifacts, which previously served as objective records of events, are now becoming potentially audience-tailored versions of reality, capable of precisely imitating individuals' behavior, appearance, and speech. As a consequence, visual and audio material ceases to be regarded as a neutral carrier of information, and its credibility can no longer be assessed solely on the basis of legal procedures or the intuitive judgment of recipients.[25]

Contemporary evidentiary standards therefore require the incorporation of advanced technical verification methods, including the analysis of manipulation traces,

21 Shoaib, Mohamed & Wang, Zefan & Taleby Ahvanooey, Milad & Zhao, Jun. (2023). Deepfakes, Misinformation, and Disinformation in the Era of Frontier AI, Generative AI, and Large AI Models. 10.1109/ICCA59364.2023.10401723.

22 Furizal, Alfian Ma'arif, Hari Maghfiroh, Iswanto Suwarno, Denis Prayogi, Kariyamin, Syahrani Lonang, Abdel-Nasser Sharkawy, Social, legal, and ethical implications of AI-Generated deepfake pornography on digital platforms: a systematic literature review, Social Sciences & Humanities Open, Volume 12, 2025, 101882, ISSN 2590-2911, https://doi.org/10.1016/j.ssaho.2025.101882.

23 Malatji, M. & Tolah, Alaa. (2024). Artificial intelligence (AI) cybersecurity dimensions: a comprehensive framework for understanding adversarial and offensive AI. AI and Ethics. 5. 883-910. 10.1007/s43681-024-00427-4.

24 Brewster, T. (2021). Fraudsters cloned company director's voice in $35 million bank heist, police find. Forbes

25 Syed Ali Hussain, Hala Georges, Nawal Abdel Razaq Askar, Illusion or reality: Ethical navigation in times of AI-enhanced visual communication, Social Sciences & Humanities Open, Volume 12, 2025, 102228, ISSN 2590-2911, https://doi.org/10.1016/j.ssaho.2025.102228.

detection of generative artifacts, and assessment of metadata consistency. This implies a shift in the evaluation of evidence from the perceptual level to the expert level, in which analytical tools, competencies in digital forensics, and interdisciplinary knowledge concerning the functioning of generative models play a decisive role. Consequently, the redefinition of digital evidence becomes not only a technological challenge but also an epistemological and legal one, transforming the manner in which state institutions, law enforcement agencies, and courts understand, interpret, and assess the authenticity of presented materials.[26]

As a result, the importance of institutional and technological mechanisms for information validation is increasing. State authorities must develop capabilities for the rapid analysis of the authenticity of digital materials, leading to the integration of artificial intelligence tools into investigative processes. Paradoxically, this means that systems used to generate threats simultaneously become indispensable elements of their detection.

At the same time, an increasing need can be observed to strengthen societal resilience against informational manipulation. Under conditions in which synthetic content may be generated on a mass scale and tailored to the recipient's profile, media literacy and the ability to critically analyze information become of fundamental importance. Without this component, even the most advanced detection systems remain insufficient in relation to the scale of influence exerted by information operations.

Another area of transformation concerns the integration of artificial intelligence into the structures of institutions responsible for internal security. AI is becoming a tool supporting the analysis of operational data, the identification of criminal patterns, and the forecasting of potential threats. In practice, this signifies a transition from a reactive operational model, based on the analysis of incidents that have already occurred, to a predictive model in which the early detection of anomalies and warning signals plays a crucial role. Simultaneously, this process generates new regulatory challenges. The development of AI-based systems requires the adaptation of legal frameworks to a dynamically changing technological reality, particularly in the areas of data processing, automated decision-making, and liability for the actions of algorithmic systems. In this context, maintaining a balance between operational effectiveness and the protection of individual rights – including the right to privacy and control over personal data – becomes of fundamental importance.[27]

At the strategic level, an important element of the state's response to emerging threats is also the development of information early warning systems. Their purpose is to identify disinformation campaigns, detect coordinated activities within

26 Korshunov, Pavel & Marcel, Sebastien. (2022). Improving Generalization of Deepfake Detection With Data Farming and Few-Shot Learning. IEEE Transactions on Biometrics, Behavior, and Identity Science. PP. 1-1. 10.1109/TBIOM.2022.3143404.

27 José Manuel Marcos-Vílchez, Milagrosa Sánchez-Martín, José Antonio Muñiz-Velázquez, Effectiveness of training actions aimed at improving critical thinking in the face of disinformation: a systematic review protocol, Thinking Skills and Creativity, Volume 51, 2024, 101474, ISSN 1871-1871, https://doi.org/10.1016/j.tsc.2024.101474.

the digital environment, and analyze potential threats to institutional and social stability. These systems are increasingly based on the integration of data from multiple sources and their processing in near real time.

All these processes indicate that internal security under conditions of artificial intelligence dominance ceases to be an exclusively reactive domain and instead becomes a complex system of continuous adaptation. The state operates within an environment in which threats are dynamic, generated, and modified in real time, thereby necessitating an equally dynamic institutional response. Consequently, security is no longer a condition that can be permanently achieved, but rather a process of maintaining relative equilibrium under conditions of permanent variability within the information environment.[28]

Conclusion

The analysis presented in this study leads to the conclusion that artificial intelligence does not merely constitute another stage in the evolution of digital tools, but rather a factor that structurally transforms the nature of contemporary threats in the field of internal security. Its significance derives not only from the automation of technical processes, but above all from its influence on the generation, processing, and credibility of information. Contemporary cybercrime and influence operations are no longer merely collections of separate activities with clearly identifiable perpetrators and objectives. Increasingly, they assume the form of complex, adaptive systems in which the boundary between human activity and algorithmic action becomes blurred. In this sense, the key transformation lies not solely in the increased scale of threats, but in their qualitative transformation – a shift from incidental activities to continuous, automated, and difficult-to-attribute processes.

A particularly significant consequence of this transformation is the weakening of traditional reference points for assessing the credibility of information. The development of generative technologies, including deepfake materials, has led to a situation in which image and sound no longer serve as stable evidence of reality. Consequently, the focus of security shifts from the protection of information itself to the protection of the capacity to interpret it.

At the same time, the observed phenomena indicate that artificial intelligence operates within a bidirectional environment. On the one hand, it strengthens the capabilities of perpetrators by enabling the automation, personalization, and scaling of criminal activities. On the other hand, it becomes an indispensable component of defense systems, utilized for threat detection, data analysis, and risk forecasting. This duality gives rise to a permanent technological tension that may be described as a persistent state of algorithmic competition.

28 Soyele, Adesola. (2025). Cross platform anomaly detection using hybrid AI models for multi-layered financial fraud in decentralized systems. International Research Journal of Modernization in Engineering Technology and Science. 07. 7. 10.56726/IRJMETS70529.

In this context, the conclusion that internal security can no longer be understood as a condition of achievable stability becomes particularly significant. It instead becomes a process of continuous adaptation in which state institutions operate under conditions of an constantly evolving information environment. This necessitates a transition from a reactive model to a predictive model based on the early detection of threats and the integration of AI-based analytical tools. Simultaneously, the development of these technologies reveals significant legal and ethical tensions. In particular, these concern the limits of permissible interference with individual privacy, the automation of administrative decision-making, and liability for the actions of algorithmic systems. Maintaining a balance between the effectiveness of state action and the protection of fundamental rights becomes one of the priority challenges of contemporary security systems.

Ultimately, it may be concluded that artificial intelligence not only transforms the tools used in cyberspace, but also redefines the very nature of risk itself. It introduces the security environment into a state of permanent variability in which advantage is gained not by those possessing the greatest resources, but by those capable of adapting most rapidly to change and integrating technological, operational, and institutional knowledge. In this sense, the key challenge for internal security lies not solely in counteracting threats generated by artificial intelligence, but in the capacity to function within a reality in which the very nature of the threat is dynamic, multilayered, and co-created by the very technologies simultaneously employed to combat it.

Bibliography

Awotidebe, M. (2023). *The Rise of Intelligent Threats: Exploring AIDriven Cybercrime in the Digital Era*.

Bada, M., Nurse, J. R. C. (2019). Developing cybersecurity education and awareness programmes for small and mediumsized enterprises (SMEs). *Information and Computer Security, 27*(3), 393–410. https://doi.org/10.1108/ICS-07-2018-0080

Brewster, T. (2021). Fraudsters cloned company director's voice in $35 million bank heist, police find. *Forbes*.

Chang, L. Y.C., Zhou, Y., Phan, D. H. (2023). Virtual kidnappings: Online scams with "Asian characteristics" during the pandemic. In R. G. Smith et al. (Eds.), *Cybercrime in the pandemic, the digital age and beyond*. Palgrave Macmillan. https://doi.org/10.1007/978-3-031-29107-4_6

Chesney, R., Citron, D. K. (2019). Deep fakes: a looming challenge for privacy, democracy, and national security. *California Law Review, 107*, 1753–1820.

Edwards, B., Hofmeyr, S., Forrest, S. (2015). Hype and heavy tails: a closer look at data breaches.

ENISA. (2023). *ENISA Threat Landscape 2023*. European Union Agency for Cybersecurity.

Europol. (2017). *Online sexual coercion and extortion as a form of crime affecting children: Law enforcement perspective* (updated 2021). Europol.

Furizal, A. M., Maghfiroh, H., Suwarno, I., Prayogi, D., Kariyamin, S., Lonang, S., & Sharkawy, A.N. (2025). Social, legal, and ethical implications of AIgenerated deepfake pornography on digital platforms: a systematic literature review. *Social Sciences & Humanities Open, 12*, 101882. https://doi.org/10.1016/j.ssaho.2025.101882

Gibson, P., Connolly, S. (2023). Fit for purpose? Taking a closer look at the UK's Online Media Literacy Strategy. *Journal of Media Literacy Education, 15*(1), 109–115.

Hong, J. (2012). The state of phishing attacks. *Communications of the ACM, 55*(1), 74–81. https://doi.org/10.1145/2063176.2063197

Jimmy, F. (2023). Understanding ransomware attacks: Trends and prevention strategies. *Journal of Knowledge Learning and Science Technology, 2*, 180–210. https://doi.org/10.60087/jklst.vol2.n1.p214

Kolla, J., Shinde, P., Baig, M., & Reddy, K. (2022). a comparison study of machine learning techniques for phishing detection. *Journal of Business and Information Systems, 4*(1), 21–33. https://doi.org/10.36067/jbis.v4i1.120

Kolias, C., Kambourakis, G., Stavrou, A., Voas, J. (2017). DDoS in the IoT: Mirai and other botnets. *Computer, 50*(7), 80–84.

Korshunov, P., Marcel, S. (2022). Improving generalization of deepfake detection with data farming and fewshot learning. *IEEE Transactions on Biometrics, Behavior, and Identity Science*. https://doi.org/10.1109/TBIOM.2022.3143404

Kshetri, N. (2021). *Cybersecurity management: An organizational and strategic approach.* https://doi.org/10.3138/9781487531249

Mahboubi, A., Luong, K., Aboutorab, H., Bui, H. T., Camtepe, S., Ansari, K., Barry, B. I. A. (2025). The evolving botnet threat landscape: a comprehensive analysis of detection techniques in the age of artificial intelligence. *Internet of Things, 33*, 101728. https://doi.org/10.1016/j.iot.2025.101728

Malatji, M., Tolah, A. (2024). Artificial intelligence (AI) cybersecurity dimensions: a comprehensive framework for understanding adversarial and offensive AI. *AI and Ethics, 5*, 883–910. https://doi.org/10.1007/s43681-024-00427-4

MarcosVílchez, J. M., SánchezMartín, M., & MuñizVelázquez, J. A. (2024). Effectiveness of training actions aimed at improving critical thinking in the face of disinformation: a systematic review protocol. *Thinking Skills and Creativity, 51*, 101474. https://doi.org/10.1016/j.tsc.2024.101474

Pum, M. (2022). *The rise of AIpowered cybercrime: Implications for digital security and policy.*

Romanosky, S. (2016). Examining the costs and causes of cyber incidents. *Journal of Cybersecurity, 2*(2), tyw001. https://doi.org/10.1093/cybsec/tyw001

Schmitt, M., Flechais, I. (2024). Digital deception: Generative artificial intelligence in social engineering and phishing. *Artificial Intelligence Review, 57.* https://doi.org/10.1007/s10462-024-10973-2

Singh, S., Dhumane, A. (2025). Unmasking digital deceptions: An integrative review of deepfake detection, multimedia forensics, and cybersecurity challenges. *MethodsX, 15*, 103632. https://doi.org/10.1016/j.mex.2025.103632

Shoaib, M. R., Wang, Z., Ahvanooey, M. T., Zhao, J. (2023). Deepfakes, misinformation, and disinformation in the era of frontier AI, generative AI, and large AI models. In *2023 International Conference on Computer and Applications (ICCA)* (pp. 1–7). IEEE.

Soyele, A. (2025). Crossplatform anomaly detection using hybrid AI models for multilayered financial fraud in decentralized systems. *International Research Journal of Modernization in Engineering Technology and Science, 7*(7). https://doi.org/10.56726/IRJMETS70529

Syed Ali Hussain, S., Georges, H., Askar, N. A. R. (2025). Illusion or reality: Ethical navigation in times of AIenhanced visual communication. *Social Sciences & Humanities Open, 12*, 102228. https://doi.org/10.1016/j.ssaho.2025.102228

Whyte, C. (2020). Deepfake news: AIenabled disinformation as a multilevel public policy challenge. *Journal of Cyber Policy, 5*(2), 199–217. https://doi.org/10.1080/23738871.2020.1805482

About the Author

Dominika Grzybowska-Ganszczyk, PhD – is a physiotherapist, medical rescue specialist, and Assistant Professor at the Academy of Physical Education in Katowice, specializing in neurorehabilitation, cardiac rehabilitation, and tactical medicine. Her scientific work includes interdisciplinary research projects conducted in collaboration with biomedical engineering and biomechatronics teams, as well as studies on functional performance and human resilience under physical and environmental stress. She integrates clinical, academic, and instructional expertise—covering self-defense, intervention tactics, and functional training—to support initiatives relevant to the health and operational safety of uniformed services and the broader society.