Professional paper

THE ROLE OF AI IN INFORMATION AND CYBER SECURITY MANAGEMENT

Tomić Rotim, S.[1], **Kutnjak, J.**[2]

[1] Zavod za informatičku djelatnost Hrvatske, Zagreb,
University of Applied Sciences Velika Gorica,
Velika Gorica, stomic@zih.hr

[2] University of Applied Sciences Velika Gorica,
Velika Gorica, josip.kutnjak1@vvg.hr

Abstract: *This paper explores the transformative role of Artificial Intelligence (AI) in information and cybersecurity management, offering a comprehensive review of current research, practical applications, and future perspectives. By analyzing recent literature and evaluating real-world case studies across sectors such as finance, telecommunications, and healthcare, the study highlights the advantages of AI in threat detection, risk assessment, fraud prevention, and compliance with security standards like ISO/IEC 27001 and the NIS 2 Directive. The research emphasizes the efficiency of AI-driven systems in identifying sophisticated cyber threats, automating responses, and improving the effectiveness of security frameworks. An iterative five-phase implementation model is also presented, along with comparative performance results of AI algorithms, demonstrating their practical value. The findings underscore AI's growing impact and provide valuable insights for organizations aiming to enhance their cybersecurity posture through intelligent and adaptive technologies.*

Keywords: *Artificial Intelligence, Information Security, Cybersecurity, Machine Learning, Threat Detection*

1. INTRODUCTION

In an era of rapid digital transformation, information security has become a fundamental concern for corporations, governmental institutions, and private entities. The exponential growth of data, the expansion of the Internet of Things (IoT), and the increasing complexity of cyber threats necessitate advanced security mechanisms. Artificial Intelligence (AI) is playing an increasingly crucial role in identifying, preventing, and mitigating cyber threats by leveraging machine learning and deep learning techniques to detect anomalies and automate security responses.

This paper explores the integration of AI into information security management, focusing on threat detection, automated response mechanisms, security policy adaptation, and compliance with global security standards such as ISO/IEC 27001. AI-driven systems enhance cybersecurity strategies by predicting potential attacks, identifying vulnerabilities, and enabling real-time responses to security incidents. The study highlights AI's contributions to

risk assessment, security protocol optimization, and the implementation of intelligent defense mechanisms.
Through different case studies, the paper illustrates how AI is revolutionizing cybersecurity frameworks, offering organizations proactive and adaptive security solutions. As AI continues to evolve, its role in information security is expected to expand, shaping the future of autonomous security systems capable of independent threat assessment and mitigation.

2. PROBLEM ANALYSIS

The rapid advancement of technology has led to significant improvements and greater success for companies, but it has also inevitably resulted in increasingly sophisticated cyber attacks. These attacks have become more effective in achieving their goals, whether financial, aimed at disrupting the operations of specific entities, or related to industrial espionage (Cucu et al., 2019). In recent years, many authors have conducted literature reviews focused on the application of artificial intelligence (AI) in cybersecurity and have provided guidelines for its implementation. For example, Jada and Mayayise analyzed 73 academic papers published between 2018 and 2023 concerning the use of AI in this field. They concluded that the advantages of AI in cybersecurity are still not sufficiently defined. According to their findings, the key benefits of AI include task automation, threat detection, and vulnerability management. However, they also highlight several challenges associated with AI, such as AI-driven attacks, lack of high-quality data, shortage of skilled professionals, and the cost of necessary resources and infrastructure (Jada and Mayayise, 2024). They emphasize the need for further research to optimize the implementation and regulation of AI solutions, especially across different sectors and types of organizations.

In his research, Mughal explores the role of artificial intelligence in the field of information security. The author identifies key advantages such as the automation of defense activities, rapid data analysis, learning from previous attacks, and using this knowledge to enhance security systems (Mughal, 2018). He outlines several important future research directions, including the development of AI systems whose decisions can be clearly understood and explained to users, the advancement of hybrid systems that combine AI with traditional methods, improving the security of AI itself, as well as standardization and regulatory development in this area. This has proven to be accurate, as the European Union has continued to develop regulations in this domain (EU Regulation, 2024). AI represents a powerful tool in the fight against cyber threats due to its ability to learn quickly, its scalability, and its capacity for automation. However, its implementation must be carefully planned, taking into account potential risks, limitations, and ethical challenges.

The literature also examines key challenges in the application of artificial intelligence in the field of cybersecurity and offers possible approaches to classification based on types of threats, attack models, and defensive techniques (Hashmi et al., 2024). Due to their ability to process

large volumes of data and recognize patterns that precede attacks, AI systems enable rapid and reliable responses to increasingly sophisticated threats. The authors propose a taxonomy based on three key aspects: threats and attack vectors, AI-based defense mechanisms, and the challenges of implementing AI in security systems. It is essential to establish strong collaboration between domain experts and AI specialists, as this is a crucial prerequisite for developing models that are both functional and robust.

An analysis of the available literature shows that in the past five years there has been a significant increase in publications on the application of artificial intelligence in the field of cybersecurity. This trend places the topic at the center of attention for both researchers and practitioners.
Through topic modeling, seven main thematic areas have been identified (Achuthan et al., 2024):

- Malware detection
- Threat and anomaly detection
- User behavioral analysis
- AI techniques in authentication
- Privacy and data processing
- Attacks against AI systems
- Explainable Artificial Intelligence (XAI)

According to Achuthan (Achuthan et al., 2024), the largest number of publications comes from China, the United States, and India, primarily from their universities or university-affiliated research networks. Despite this, it can still be said that the development of explainable and ethically responsible AI systems has not yet been fully established. Much work remains to ensure that AI is applied in a transparent, functional, and ethical manner across all aspects of cybersecurity.

Other authors also identify key areas for the future development of artificial intelligence, including the creation of explainable AI (XAI) models, the use of hybrid approaches that combine AI with traditional security methods, and the standardization and regulation of AI technologies in security environments (Abbas et al., 2019). A combination of technical advancement and ethical responsibility is essential for the safe and effective application of AI in defending digital systems.

For the successful implementation of AI in the field of cybersecurity within a corporate environment, a five-phase iterative process model can be applied (Lier et al., 2025). This model consists of five iterative phases and 19 steps:

1. Requirements Analysis – defining business and security needs.
2. Strategic Analysis – evaluating internal and external aspects of building the AI solution, selecting the mode of operation (detection only or also autonomous defense), and testing in a sandbox environment.

3. Testing & Learning – training the model, evaluating it, optimizing, and repeating the process as needed.
4. Implementation – integrating the solution into the IT infrastructure, involving stakeholders, and launching pilot projects.
5. Evaluation – continuous performance monitoring, assessment based on predefined KPIs, adaptation, and iteration.

To measure the effectiveness of the model, seven key indicators have been defined, such as the false positive/negative alert rate, energy consumption cost, system efficiency, and others. This model offers a solid framework for the development and application of AI systems in the field of cybersecurity. However, it definitely requires further research and validation of its effectiveness and reliability, as well as continuous improvement.

A particularly important role of artificial intelligence and data mining techniques is recognized in the context of developing threat detection systems - Cyber Threat Intelligence (Gupta et al., 2019). In today's digital environment, where organizations handle vast amounts of data and face increasingly sophisticated attacks, traditional defense methods have proven insufficient. Therefore, the authors propose the integration of advanced AI models and data mining algorithms to enhance threat recognition and mitigation.

The authors systematically analyze existing techniques and algorithms, such as classification, clustering, prediction, neural networks, support vector machines, and genetic algorithms, which are used for anomaly detection and identifying threat behavior patterns. Additionally, they emphasize the importance of statistical methods like ANOVA and correlation analysis for identifying relationships among attributes in real-world datasets, such as those related to financial fraud. They stress that perimeter-based defense systems alone are no longer sufficient to cope with the complexity of modern threats. What is needed is a comprehensive integration of multi-layered AI systems capable of autonomously analyzing threats, learning from past attacks, and predicting new ones.

From all of the above, it is evident that the authors in this field have focused on developing AI models that will assist in identifying cyber threats and ensuring timely and effective protection. In addition, this paper will cover another significant area, how AI can be used in the development of information and cybersecurity management systems, particularly in the areas of risk assessment, identification of mitigation measures, and their development and implementation, which is also a requirement under the NIS 2 Directive (EU Regulation, 2022). This area has not yet been adequately addressed in existing studies, and this paper will demonstrate, through a case study, the potential application of AI in this specific domain.

3. METHODS

The methodology of this paper is based on a comprehensive review of scientific literature, case studies, and practical implementations of artificial intelligence (AI) in information security management. A qualitative research approach is adopted, focusing on theoretical analysis and real-world applications to assess the effectiveness of AI-driven security mechanisms.

- Literature Review – The study begins with an extensive analysis of existing research, academic papers, and industry reports related to AI in cybersecurity. This review identifies key trends, challenges, and opportunities in the field.
- Comparative Analysis – Different AI techniques, including machine learning, deep learning, and automated security protocols, are compared based on their efficiency in identifying, preventing, and mitigating cyber threats.
- Case Studies – The research includes real-world examples from industries such as finance, healthcare, and telecommunications to illustrate how AI is transforming information and cyber security.
- ISO/IEC 27001 Compliance Analysis – A detailed examination of how AI can support compliance with international security standards, focusing on risk management, policy adaptation, and automated security responses.
- Future Perspectives – The study concludes with a discussion on the future role of AI in cybersecurity, emphasizing potential advancements, ethical considerations, and challenges that need to be addressed.

4. DISCUSSION AND CONCLUSION

This study aims to provide a comprehensive understanding of the role of Artificial Intelligence (AI) in information security management, identifying its benefits, challenges, and future implications. The key outcomes of this research include a deeper understanding of the application of AI in cybersecurity and the identification of effective AI-based security measures through an extensive literature review. It offers a detailed analysis of how AI technologies contribute to threat detection, attack prevention, and automated security responses, as well as insights into the efficiency of various AI-driven security mechanisms. These findings help organizations implement advanced and tailored solutions to meet their specific security needs.

In addition, the purpose of this research is to highlight the importance of AI in establishing information and cybersecurity management systems and achieving compliance with ISO/IEC 27001 (ISO/IEC, 2022) and cybersecurity regulatory frameworks (EU Regulation, 2022). The study provides a clear assessment of how AI supports organizations in aligning with

international security standards, with a particular focus on risk assessment, policy enforcement, and automated compliance monitoring.

Furthermore, the paper presents the results of several case studies and practical applications. This is particularly important for understanding real-world implementations of AI across various industries, including finance, healthcare, and telecommunications, demonstrating its impact on modern cybersecurity frameworks.

4.1. EVALUATION OF AI'S ROLE IN ISO/IEC 27001 COMPLIANCE

The effective application of Artificial Intelligence (AI) plays a crucial role in the development of information and cybersecurity management systems. Using the example of a healthcare organization required to comply with the NIS 2 Directive and to assess all potential risks related to its information assets, particularly software and hardware, the key success factor lies in identifying current risks and determining their likelihood and impact in case of materialization. Within such a healthcare organization, a list of critical information assets was compiled, focusing on confidentiality, integrity, and availability. The team responsible for the implementation then conducted a risk assessment and identified appropriate risk treatment measures, a small excerpt and illustrative example of which is presented in Table 1 from the broader risk register.

Table 1. Public Health Cybersecurity Risk Register

Risk	*Likelihood*	*Impact*	*Risk Mitigation Measurement*
Unauthorized access to health records	High	Critical	Implement multi-factor authentication and audit logging
Phishing attacks targeting staff	High	High	Conduct regular phishing simulations and training
Malware infection through email	Medium	High	Use advanced anti-malware tools and sandboxing
Data breach via third-party vendor	Medium	Critical	Vendor risk assessment and regular audits
Ransomware attack on hospital systems	High	Critical	Regular backups and incident response plan
Inadequate access controls	Medium	Medium	Enforce least privilege policy and periodic review
Use of outdated software	High	High	Regular updates and vulnerability patching
Loss of data due to backup failure	Low	High	Test and monitor backup systems regularly
Distributed Denial of service (DDoS) attack	Medium	Medium	Deploy network firewalls and intrusion detection
Insider threat from disgruntled employee	Low	High	Background checks and activity monitoring

Initially, the risk assessment was carried out by the team members without the use of artificial intelligence. Subsequently, AI was introduced into the risk identification process for validation purposes. The results demonstrated a significant advantage, 35% more relevant risks were identified with the help of AI, risks that the team had not initially recognized, and for which they had not defined appropriate mitigation measures. The application of AI significantly accelerated and improved the overall risk management process, reducing both the likelihood and potential impact of risks. This improvement was reflected in a noticeable decrease in the number of security incidents over the following year (see Figure 1).

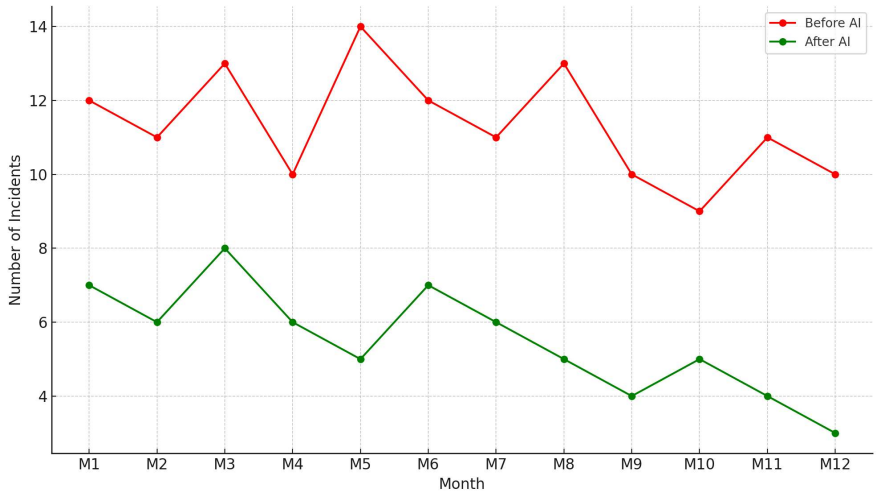

Figure 1. Monthly Security Incidents before and after the implementation of AI in risk management

4.2. CASE STUDY ANALYSIS AND PRACTICAL APPLICATIONS

One of the key outcomes of this study is the analysis of multiple case studies and the practical application of AI across various industries, including the financial sector, telecommunications, and healthcare. In the financial sector, a common focus of research is the effectiveness of different machine learning algorithms in detecting fraud in credit card transactions. As the financial industry increasingly relies on digital transactions, the rise in cyber fraud presents a significant challenge.

Table 2 provides a comparison of nine machine learning models: Random Forest, Decision Tree, K-Nearest Neighbors, Logistic Regression, Gradient Boosting, AdaBoost, Extra Trees, MLP, and Naive Bayes (Idrees et al., 2024).

Table 2: Performance of 9 AI Models for Fraud Detection

Model	*Accuracy (%)*	*Precision (%)*	*Recall (%)*	*F1 Score (%)*

Random Forest	94.9991	95.9887	95.1234	95.1102
Gradient Boosting	94.7425	94.5634	93.8789	94.2196
AdaBoost	94.5238	94.2871	92.8876	93.5721
Extra Trees	94.6895	94.1108	93.0198	93.5600
Decision Tree	93.6752	92.6543	91.3452	92.0000
K-Nearest Neighbors	93.3487	91.2087	89.6754	90.4231
Logistic Regression	93.8764	92.7653	91.4432	92.0990
Naive Bayes	92.0873	90.0987	88.2103	89.1345
MLP	94.0157	93.1011	92.4456	92.7721

The experiments were conducted on a real, publicly available dataset containing over 284,000 transactions, of which only 0.17% were labeled as fraudulent. Due to the extreme class imbalance, the challenge of detection was significantly more complex. The authors used standard evaluation metrics such as accuracy, precision, recall, F1-score, and the false positive rate. The results show that the Random Forest model achieved the highest performance, with an accuracy rate of nearly 95% and a very high F1-score, making it the most suitable model for fraud detection. The study also addressed issues such as dataset imbalance and the importance of selecting appropriate evaluation metrics when dealing with rare events.

An interesting case in the telecommunications industry analyzes the development of a model for detecting International Revenue Share Fraud (IRSF) and other types of telecom fraud (Yehya, 2023). The system is based on a comprehensive dataset consisting of Call Detail Records (CDRs) from a real telecom network. Multiple machine learning models were employed in the development process, including classifiers and anomaly detection techniques. To identify unusual patterns, models such as Random Forest, Support Vector Machine (SVM), and Isolation Forest were applied.
Additionally, a hybrid model combining a Graph Attention Network (GAT) and a Gated Recurrent Unit (GRU), integrated with Isolation Forest, was experimentally tested. This model achieved a high recall and moderate precision, indicating a strong ability to detect nearly all fraudulent activities while maintaining moderate accuracy in classification. The system demonstrated a clear advantage over traditional manual methods by offering higher sensitivity (recall), detecting almost all fraud cases, and maintaining a well-balanced trade-off with precision. The model was tested on real-world data from a telecom operator, confirming its practical applicability.

In the healthcare sector, multi-level systems have been developed for detecting fraud in health insurance by combining adaptive machine learning methods with deep learning techniques (Matloob et al., 2025). The system is built on a hierarchical architecture that analyzes the

behavior of patients, service providers, and pharmacies based on historical data and behavioral patterns (see Figure 2).
The authors emphasize that traditional rule-based approaches have limited capability in detecting sophisticated and newly emerging fraudulent activities. Therefore, their approach integrates neural models, including a Transformer model, for detecting anomalies in time-series data.

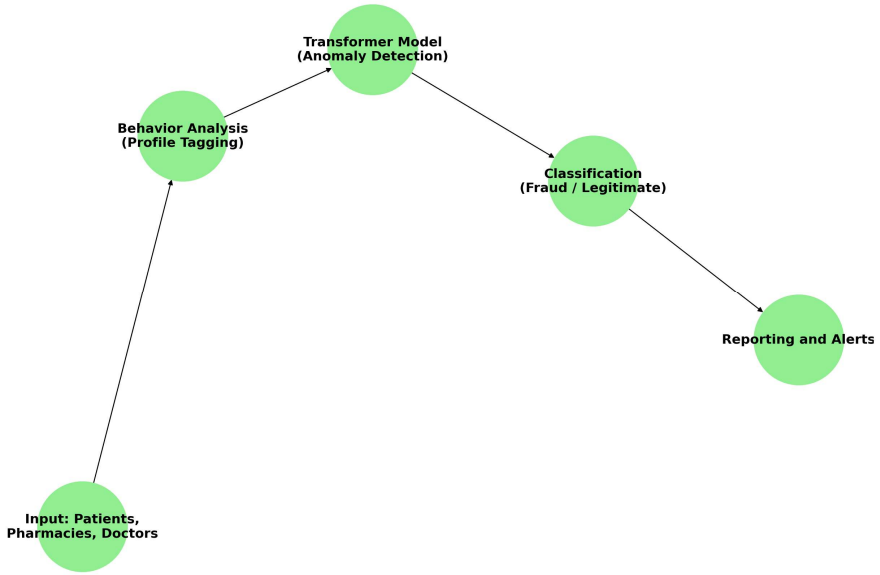

Figure 2. Schematic Diagram of the Healthcare Fraud Detection System

The system utilizes prior behavior classification to enable the model to distinguish between legitimate and fraudulent actions within the context of patients and healthcare processes. During the experimental evaluation, a real dataset was used, containing information on transactions and employee behavior within a hospital, which allowed the model's effectiveness to be tested in an authentic environment.
The model achieved an accuracy rate of 97%, significantly outperforming the effectiveness of traditional rule-based systems.

The future role of AI in cybersecurity is expected to evolve significantly, driven by continuous advancements in machine learning, deep learning, and automation technologies. AI systems will likely move beyond reactive defense mechanisms toward proactive and predictive security models capable of anticipating and neutralizing threats before they materialize. However, this evolution brings critical challenges that must be addressed, including the ethical use of AI, transparency through explainable AI (XAI), and the mitigation of risks associated with AI-driven attacks. Furthermore, ensuring data quality, developing standardized frameworks, and establishing regulatory compliance will be key to achieving reliable and responsible AI

integration. The findings of this study highlight the necessity of ongoing research and collaboration between domain experts, policymakers, and technologists to harness AI's full potential while minimizing associated risks.

5. REFERENCES

Abbas, N. N., Ahmed, T., Shah, S. H. U. et al. (2019). Investigating the applications of artificial intelligence in cyber security. Scientometrics, 121, 1189–1211. https://doi.org/10.1007/s11192-019-03222-9

Achuthan, K., Ramanathan, S., Srinivas, S., Raman, R. (2024). Advancing cybersecurity and privacy with artificial intelligence: current trends and future research directions, Frontiers, 05 December 2024, DOI 10.3389/fdata.2024.1497535

Cucu, C., Gavrioloaia, G., Bologa, R., & Cazacu, M. (2019). Current technologies and trends in cybersecurity and the impact of artificial intelligence (Vol. 2). eLearning & Software for Education

EU Regulation 2022/2555 of European Parliament and of the Council on measures for a high common level of cybersecurity across the Union of 14. December 2022.

EU Regulation 2024/1689 of European Parliament and of the Council laying down harmonised rules on artificial intelligence of 13 June 2024.

Gupta, S., Sabitha, S., Punhani, R. (2019). Cyber Security Threat Intelligence using Data Mining Techniques and Artificial Intelligence, International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Hashmi, E., Yamin, M. M., Yayilgan, S. Y. (2024). Securing tomorrow: a comprehensive survey on the synergy of Artificial Intelligence and information security. AI Ethics. https://doi.org/10.1007/s43681-024-00529-z

Idrees, A. M., Elhusseny, N. S., Ouf, S. (2024). Credit Card Fraud Detection Model-based Machine Learning Algorithms, IAENG International Journal of Computer Science, Volume 51, Issue 10, October 2024, Pages 1649-1662

ISO/IEC 27001:2022 Information security, cybersecurity and privacy protection — Information security management systems — Requirements

Jada I., Mayayise, T. O. (2024). The impact of artificial intelligence on organisational cyber security: An outcome of a systematic literature review. Data and Information Management, 8 (2), 100063. https://doi.org/10.1016/j.dim.2023.100063

Lier, S. K., Eppers, T. M., Gerlach, J., Müller, P., Breitner, M. H. (2025). An iterative five-phase process model to successfully implement AI for cybersecurity in a corporate environment, Electronic Markets (2025) 35:56, 21 May 2025., https://doi.org/10.1007/s12525-025-00802-x

Matloob, I., Khan, S., Rukaiya , R., Alfrahi, H., Khan, J. A. (2025). Healthcare fraud detection using adaptive learning and deep learning techniques, Evolving Systems 16:72, 9 May 2025, https://doi.org/10.1007/s12530-025-09698-6

Mughal, A. A. (2018). Artificial Intelligence in Information Security: Exploring the Advantages, Challenges, and Future Directions. Journal of Artificial Intelligence and Machine Learning in Management, 2(1), 22–34. https://journals.sagescience.org/index.php/jamm/article/view/51

Yehya, B. A., Salhab, N. (2023). Telecommunications Fraud Machine Learning-based Detection, October 2023, https://www.researchgate.net/publication/383177588