
The Role of Artificial Intelligence in Cybersecurity Management in the Public Sector

Submitted 13/02/26, 1st revision 05/03/26, 2nd revision 20/04/26, accepted 03/05/26

Tomasz Grudniewski¹, Marta Chodyka², Monika Ostrowska³, Zbigniew Ciekankowski⁴, Leszek Elak⁵, Valeriy Kuznetsov⁶, Zenon Zamiar⁷

Abstract:

Purpose: The aim of this article is to analyse the role of artificial intelligence (AI) in information security management in the public sector, with particular emphasis on the scope of its implementation, the identification of key risks, and the assessment of its impact on the operational efficiency of public organisations.

Design/Methodology/Approach: The study utilised an analysis of academic literature and European Union strategic and regulatory documents, including the Artificial Intelligence Act and the NIS2 Directive. Secondary data from reports by cybersecurity institutions, such as ENISA and CERT- -Polska, were also used. The research problem was formulated as follows: how does the implementation of artificial intelligence affect the effectiveness of information security management in the public sector, and what risks accompany this process? The research hypothesis adopted was that the use of artificial intelligence significantly increases the effectiveness of information security management in public organisations, but requires the simultaneous implementation of advanced risk management mechanisms and regulatory oversight.

Findings: The results of the analysis indicate that artificial intelligence significantly increases the effectiveness of information security management by improving threat detection capabilities, reducing incident response times and supporting decision-making processes. At the same time, significant risks associated with its use have been identified, including a lack

¹John Paul II University in Biala Podlaska, Poland, ORCID: 0000-0003-3394-8992, e-mail: t.grudniewski@dyd.akademiabialska.pl;

²John Paul II University in Biala Podlaska, Poland, ORCID: 0000-0002-8819-2451, e-mail: m.chodyka@dyd.akademiabialska.pl;

³A.F. Modrzewski University in Kraków, Poland, ORCID: 0000-0002-0397-1131, e-mail: m.ostrowska@onet.pl;

⁴John Paul II University in Biala Podlaska, Poland, ORCID 0000-0002-0549-894X, e-mail: zbigniew@ciekanowski.pl;

⁵War Studies University, Poland, ORCID: 0000-0002-5255-9768, e-mail: l.elak@akademia.mil.pl;

⁶Railway Research Institute, Poland, ORCID: 0000-0003-4165-1056, e-mail: vkuznetsov@ikolej.pl;

⁷International University of Transport and Logistics in Wrocław, Poland, ORCID: 0000-0001-9887-0183, e-mail: z.zamiar@wp.pl;

of algorithm transparency, the risk of decision-making errors, vulnerability to attacks, and threats to data privacy.

Practical implications: *This article contributes to the development of research on information security management by proposing an integrated security management model utilising AI that covers the strategic, tactical, and operational levels.*

Originality/value: *The results may serve as a basis for further research and support for management practice in public administration.*

Keywords: *AI Act artificial intelligence, information security management, public sector, cybersecurity, AI governance, NIS2, AI Act.*

JEL Codes: *O33, L86, H11, H12, D83.*

Paper type: *Research article.*

1. Introduction

The dynamic development of digital technologies, particularly artificial intelligence (AI), is leading to significant changes in how public administration functions and information security is managed. In an era of advancing digitalisation of public services, a growing number of cyber threats and increased reliance of state institutions on IT systems, ensuring an adequate level of information protection is becoming one of the key challenges of modern public administration.

Artificial intelligence, as a tool for automating analytical processes, threat prediction, and decision-making support, is becoming increasingly important in information security systems.

In the public sector, AI-based solutions are used across a wide range of applications, including the detection of security incidents, the analysis of network traffic anomalies, the protection of personal data, and support for crisis management processes. At the same time, the use of artificial intelligence generates new risks, including transparency of algorithms, susceptibility to manipulation, decision-making errors, and issues of accountability for actions taken. Consequently, there is a need for a comprehensive approach to information security management that takes into account both the potential and limitations of AI technology.

An important context for the problem under analysis is the regulatory changes taking place at European Union level, including in particular the implementation of the Artificial Intelligence Act and the NIS2 Directive, which impose new obligations on public institutions regarding risk management, cybersecurity and the responsible use of digital technologies (Grima *et al.*, 2023).

These regulations form the foundation for building coherent information security management systems that integrate technological, organisational and legal aspects.

Despite the growing importance of artificial intelligence in information security, there remains a research gap regarding the extent of AI implementation in the public sector and its impact on the effectiveness of security management. In particular, there is a lack of in-depth analyses that take into account both the operational benefits and the risks associated with implementing these technologies in public administration institutions.

The research perspective adopted forms part of the body of work examining security management in the context of digital transformation and the growing role of smart technologies in public administration systems. This article attempts to integrate technological, organisational and regulatory approaches, which is particularly important in the context of contemporary challenges to national security and the information society.

2. Digital Transformation and the Significance of Artificial Intelligence in Information Security

The dynamic development of digital technologies is one of the key factors determining the contemporary functioning of public administration and information security systems (Żurawski *et al.*, 2025). In particular, the growing importance of artificial intelligence (AI) influences the way information processing is managed, threat analysis, and decision-making in public institutions. The ongoing digitisation of public services, the development of e-government and the integration of ICT systems mean that information security is becoming a fundamental element of the stability of the state and its organisational structures.

Artificial intelligence enables automating data processing, analysing large datasets (big data), and identifying patterns and anomalies that may indicate potential security threats. The literature on the subject emphasises that AI significantly increases the efficiency of information management by accelerating decision-making processes and improving the quality of predictive analyses (Tveita and Hustad, 2025).

In particular, the use of machine learning algorithms enables the detection of security incidents in real time, which is crucial given the growing number of cyberattacks targeting public institutions.

The implementation of artificial intelligence in the public sector covers a wide range of applications, including threat detection systems, network traffic analysis, information access management and personal data protection. Research indicates that the use of AI in public administration contributes to increased operational efficiency, optimised resource utilisation, and improved quality of public services

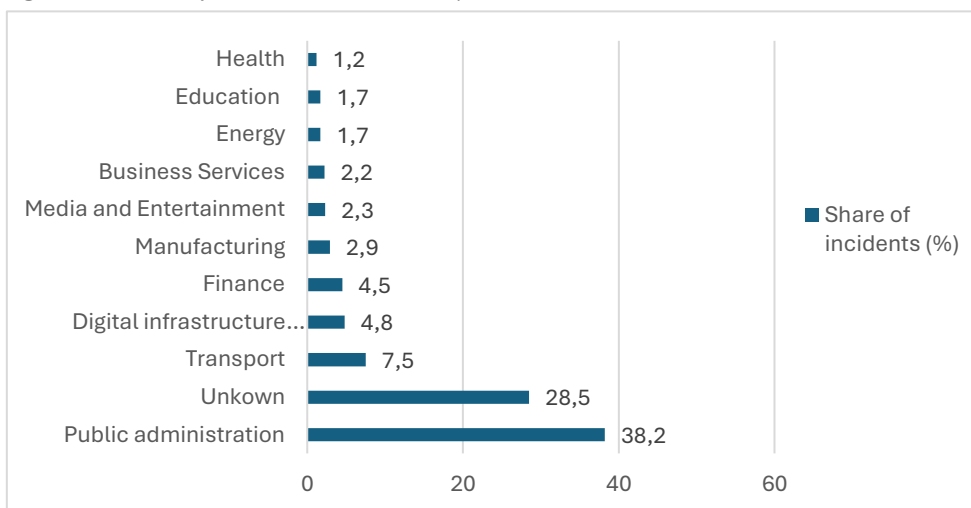
(Raksnys, Gudelis, and Guogis, 2025). Furthermore, these technologies enable the creation of more advanced systems to support crisis management and response to security incidents.

A key aspect of using AI in information security management is its ability to support decision-making processes in public administration. AI-based systems allow for the analysis of large amounts of data in a short time, which translates into greater accuracy and speed in decision-making. As research indicates, AI can significantly improve the quality of public administration by increasing the transparency of processes and reducing human error (Alhudaidi, Tomasević, and Bozić, 2025).

At the same time, the development of artificial intelligence is closely linked to the growing complexity of the digital security environment. As public administration becomes digitised, the number of potential attack vectors increases, necessitating the use of increasingly advanced information security tools. In this context, AI is becoming not only a tool supporting security management, but also a component in building the resilience of the state's information systems.

Research indicates that the integration of AI into cybersecurity systems significantly enhances the ability to detect and neutralise threats, particularly in environments with a high degree of technological complexity (Berko, 2025). Figure 1 below shows the share of recorded incidents by sector.

Figure 1. Share of recorded incidents by sector.



Source: ENISA dataset.

The rise in the number of cybersecurity incidents justifies the need to utilise AI tools. It is worth emphasising that the implementation of artificial intelligence in the public sector is an inevitable process, driven by the need to increase the efficiency of

public administration and to adapt to a rapidly changing threat landscape. As scientific analyses indicate, the development of AI in public administration requires appropriate organisational, legal and technological preparation to ensure the secure and responsible use of this technology (Kaczmarek, Karpiuk, and Spaziani, 2024).

Artificial intelligence is playing an increasingly significant role in public-sector information security management, serving as a tool to enhance the efficiency, resilience, and adaptability of public administration systems. At the same time, its development requires a comprehensive approach that balances its technological potential with the need to ensure appropriate control and oversight mechanisms.

3. Risks and Threats Associated with the Use of Artificial Intelligence in Information Security Management

The implementation of artificial intelligence in the public sector, despite numerous operational benefits, gives rise to a new category of technological, organisational and legal risks. The literature emphasises that AI can enhance the effectiveness of incident detection, automate data analysis and support decision-making processes, but at the same time generates threats arising from the opacity of models, susceptibility to errors, dependence on data quality, and difficulties in attributing responsibility for the consequences of the system's operation. In public administration, these risks carry particular weight, as they concern systems handling sensitive data, decisions affecting citizens' rights, and critical infrastructure (Okdem and Okdem, 2024).

One of the key risks is the so-called 'black box' problem, i.e. the limited explainability of AI systems. In practice, this means that a public body may use a tool that generates recommendations or classifications, but is unable to fully reconstruct the logic behind the decision made.

From an information security management perspective, this leads to reduced auditability, hinders the assessment of the system's correct operation, and may lower the level of trust in administrative processes. Research on the transparency and accountability of AI indicates that a lack of interpretability hinders both internal oversight and external scrutiny of the legality and proportionality of the solutions employed (Genaro-Moya, Lopez-Hernández, and Godz, 2025).

The second key area of risk is algorithmic bias, most often resulting from the quality of training data, design flaws or inadequate model assumptions. In the public sector, this problem can lead to unequal treatment of citizens, incorrect prioritisation of threats, and disproportionate profiling of specific groups.

The literature indicates that even technically advanced AI systems can reproduce existing biases present in the input data, and consequently exacerbate discrimination and reputational risks. In information security management, this means the need for

constant data validation, monitoring of model performance, and the implementation of procedures to correct decision-making errors (Hofmann, 2025).

Threats to privacy and data protection are also particularly significant. AI systems used in public administration often operate on large datasets, encompassing personal, behavioural and operational information. This increases the risk of unauthorised data disclosure, the secondary use of information beyond the original processing purpose, and the development of overly intrusive analytical models.

Research on AI security and privacy highlights threats such as model inversion, membership inference, data leakage and backdoor attacks, which can lead to the reconstruction or disclosure of sensitive input data. For the public sector, this means the need to closely integrate AI implementations with the principles of privacy by design, data minimisation, and strong oversight of the model lifecycle (Jada and Mayayise, 2024).

Another group of threats consists of attacks directed specifically at AI models and their operational environment. These primarily include adversarial attacks, training data poisoning, model stealing, manipulation of input data, and disruption of the learning process. From a cybersecurity perspective, this problem is particularly serious, as an AI system can simultaneously strengthen an organisation's defences and become a new attack surface.

Literature reviews indicate that the resilience of models to manipulation is becoming one of the most significant challenges for the security of modern machine learning-based systems. In public administration, the consequences of such an attack may include the misclassification of incidents, false alarms, or the failure to detect a real threat (Abomakhelb, Jalil, Buja, Alhammadi, and Alenezi, 2025).

Technological risks overlap with organisational issues. Many public institutions do not yet have mature procedures for managing the lifecycle of AI systems, adequate staff, or mechanisms for integrating new tools into the existing information security management system. As a result, the implementation of AI may lead to a phenomenon of 'pseudo-automation' of security, in which an organisation increases its dependence on technology without building the necessary competence, control and audit infrastructure.

Research on the integration of AI in public administration indicates that issues of scalability, governance, accountability and a lack of implementation standards are among the most frequently identified barriers (Aarab, El Marzouki, Boubker, and El Moutaqi, 2025).

In this context, the AI Act – EU Regulation 2024/1689, which came into force on 1 August 2024 and introduced a uniform legal framework for AI systems in the European Union – is of key importance. This Act adopts a risk-based approach: it

prohibits practices deemed unacceptable, establishes specific obligations for high-risk systems, imposes transparency requirements for selected applications, and regulates general-purpose models (Regulation (EU) 2024/1689).

From the public sector's perspective, the prohibitions on certain manipulative practices, social scoring and certain biometric applications are particularly important, as are the obligations relating to human oversight, data quality, documentation, event logging, resilience and the monitoring of high-risk systems. Some provisions came into force on 2 February 2025, others on 2 August 2025, the bulk of the regulations on 2 August 2026, and selected obligations for certain high-risk systems on 2 August 2027.

For information security management, the AI Act is significant not only in regulatory terms but also in organisational terms. The regulation forces a shift from the ad hoc implementation of AI tools to a compliance management model, in which risk assessment, documentation of system operations, ensuring human oversight, monitoring data quality and procedures for responding to anomalies become essential. This means that public institutions implementing AI should treat compliance with the AI Act as an integral part of their information security governance system, rather than merely as an external legal obligation.

Equally important is the NIS2 Directive, i.e., EU Directive 2022/2555, which entered into force on 16 January 2023, with Member States required to transpose it by 17 October 2024. NIS2 has expanded the scope of the EU cybersecurity regime in terms of both entities and subject matter, and has imposed obligations on critical and important entities regarding cyber risk management.

These include, amongst others, risk analysis, incident handling, business continuity, supply chain security, security of system procurement and development, assessment of the effectiveness of protective measures, basic cyber hygiene practices, and training (Directive (EU) 2022/2555). The Directive also strengthens requirements for incident reporting and management accountability. In practice, this creates a framework within which AI systems used for information protection must be managed as part of an organisation's broader cybersecurity system.

The relationship between the AI Act and NIS2 is complementary. The AI Act focuses on the security, transparency and compliance of the AI systems themselves, whilst NIS2 strengthens the resilience of organisations and their cybersecurity processes.

For the public sector, this means the need to build a dual governance model: on the one hand, ensuring that AI technologies comply with EU legal requirements, and on the other, ensuring the organisational resilience of the infrastructure, data and processes in which this AI operates. Only by combining these two perspectives can the risks associated with automated decision-making, the vulnerability of models to

attacks and the administration's growing dependence on intelligent systems be mitigated.

Consequently, it must be accepted that the greatest threat is not the use of artificial intelligence itself, but its implementation without appropriate managerial, technical and legal frameworks (Cheong, 2024).

In public institutions, AI can genuinely enhance information security, but only if it is accompanied by mechanisms for explainability, auditing, model resilience, data protection, human oversight and regulatory compliance. Without these conditions, the technology that was intended to mitigate risk may become an additional source of it.

4. AI Implementations in the Public Sector and Information Security Management Models

The development of artificial intelligence in the public sector is contributing to the emergence of new information security management models that integrate advanced analytical technologies with traditional management mechanisms. In administrative practice, there is a growing number of AI system implementations supporting both operational and strategic aspects of information security, which forms part of the wider process of the state's digital transformation.

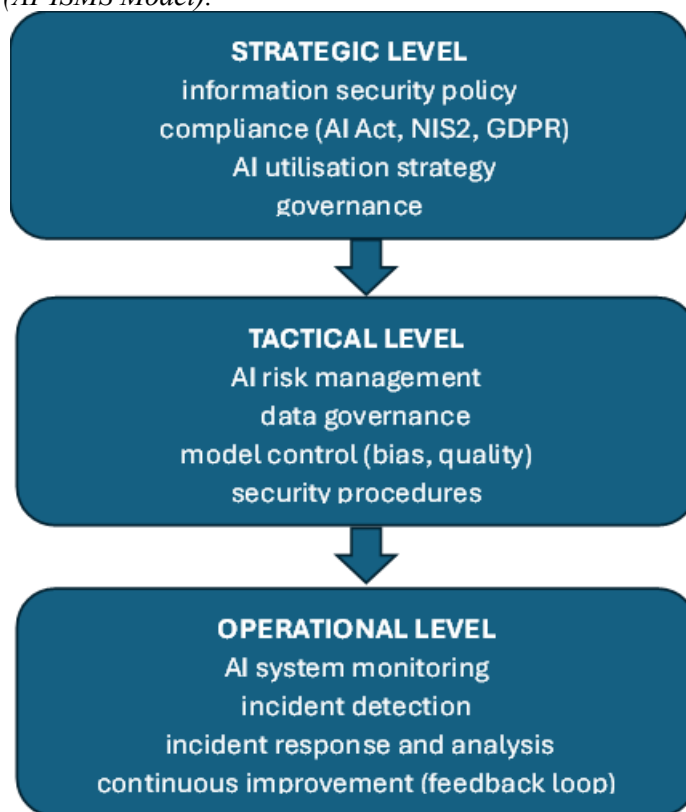
Among the most commonly identified areas of application for artificial intelligence in the public sector are cyber threat detection systems, network traffic analysis, critical infrastructure monitoring, and the automation of data protection processes. In particular, machine learning algorithms are used to identify anomalies, classify incidents and predict potential threats. Research indicates that the use of AI in cybersecurity allows for a significant increase in the effectiveness of incident detection, whilst reducing an organisation's response time

In public administration, artificial intelligence is also used in managing access to information and in decision-support systems. Examples include solutions used for risk analysis in tax systems, the identification of financial fraud, and the support of control and audit processes. The literature emphasises that AI implementations increase the transparency of public institutions' operations and improve the efficiency of information resource utilisation (Aarab *et al.*, 2025).

Another significant area of development is the use of artificial intelligence in crisis management and public safety. AI systems can support real-time data analysis, threat forecasting and the coordination of activities by services responsible for crisis response. In the context of a growing number of hybrid threats and cyberattacks, the ability to rapidly process and interpret data is becoming a key element of effective information security management (Nascimento *et al.*, 2025).

However, implementing artificial intelligence in the public sector requires adopting appropriate management models that ensure consistency among technology, organisational processes and regulatory requirements. In response to the identified challenges and regulatory requirements, a proprietary model for integrated information security management using artificial intelligence (AI-ISMS Model) has been proposed, as shown in the figure below.

Figure 2. Integrated information security management model using artificial intelligence (AI-ISMS Model).



Source: Own work.

The model adopts a multi-level approach to information security management using artificial intelligence, integrating three core management levels: strategic, tactical and operational. The strategic level encompasses the development of security policies, ensuring regulatory compliance and defining the direction of AI use within the organisation. The tactical level focuses on risk management, data quality control and oversight of AI model operations. The operational level, in turn, covers ongoing system monitoring, incident detection and threat response.

A key element of the model is the cross-cutting layer comprising AI technologies that support all levels of management through data analysis, process automation and

threat prediction. The model also takes into account the regulatory environment, in particular the requirements arising from the AI Act and NIS2, which determine how artificial intelligence systems are implemented and supervised in the public sector.

However, the implementation of artificial intelligence in the public sector requires the adoption of appropriate governance models that ensure consistency between technology, organisational processes and regulatory requirements. The literature highlights the need to build integrated information security management systems (ISMS), extended to include components related to AI governance (Tyagi *et al.*, 2023).

This entails taking into account elements such as model lifecycle management, data quality control, algorithm oversight, audit mechanisms, and incident response procedures related to the operation of AI systems (He, Davila, Bi, Wang, and Hou, 2025). Of particular importance in this context is the integration of requirements arising from the Artificial Intelligence Act and the NIS2 Directive into information security management practice.

The models implemented should adopt a risk-based approach, encompassing both technological and organisational risk assessment. It is also crucial to ensure that AI systems comply with the principles of transparency, accountability and human oversight, which form one of the cornerstones of EU regulations.

In practice, this means adopting a multi-level approach, in which information security management using AI takes place at three fundamental levels: strategic, tactical and operational. At the strategic level, it is crucial to define security policies, regulatory compliance and long-term objectives for the use of AI.

The tactical level covers risk management, the implementation of security procedures, and the integration of AI systems with existing organisational infrastructure. The operational level, on the other hand, focuses on the ongoing monitoring of systems, incident response, and real-time data analysis.

Research indicates that the effectiveness of AI implementations in the public sector depends on the institution's level of organisational maturity, the availability of technological resources, and staff competencies.

In particular, the importance of training, building cybersecurity awareness and developing analytical skills among public administration staff is emphasised (Jada and Mayayise 2024).

The implementation of artificial intelligence in the public sector is a key element in the transformation of information security management systems. The integration of AI technologies with existing organisational structures and regulatory frameworks enables an increase in the operational efficiency of public institutions, an

improvement in the quality of analysis, and a strengthening of the resilience of information systems.

At the same time, the effectiveness of these implementations depends on the ability to build coherent management models that combine technological, organisational and legal aspects within a unified information security system.

5. Conclusion

An analysis of the role of artificial intelligence in information security management in the public sector indicates that these technologies are becoming one of the key elements of modern national security systems. Their significance stems primarily from their ability to process large datasets, identify threat patterns and support real-time decision-making processes.

In the context of a growing number of cyber threats and the ongoing digitalisation of public administration, the use of AI enables increased operational efficiency, shorter incident response times and improved information management.

At the same time, the analysis confirms that the implementation of artificial intelligence entails significant risks that may affect the level of information security. The most important of these include issues related to the transparency of algorithms, vulnerability to errors and manipulation, the risk of privacy breaches, and difficulties in assigning responsibility for decisions made by AI systems. In the public sector, where sensitive data is processed and decisions of high social significance are made, these threats require particular attention and a systemic approach to management.

A key element in mitigating these risks is the implementation of appropriate regulatory and organisational frameworks (Ciekanowski *et al.*, 2023). In this context, European Union regulations such as the Artificial Intelligence Act and the NIS2 Directive are of particular importance, as they introduce a risk-based approach, security management obligations, and requirements regarding transparency and oversight of AI systems.

These regulations form the foundation for building modern information security management systems that integrate technological, legal and organisational aspects.

In light of the analyses conducted, it must be concluded that the effective use of artificial intelligence in the public sector requires the adoption of an integrated information security management model.

This model should take into account the strategic level (defining policies and regulatory compliance), the tactical level (risk management and implementation of procedures), and the operational level (monitoring systems and responding to incidents).

Of particular importance is also the development of staff competencies, building cybersecurity awareness, and ensuring appropriate audit and control mechanisms.

The findings indicate that artificial intelligence has the potential to significantly strengthen information security systems in public administration; however, its effectiveness depends on the method of implementation and the level of organisational maturity of the institution.

The implementation of AI should not be treated solely as a technological solution, but as part of a broader security management strategy, encompassing legal, organisational and social aspects as well. Only such an approach allows the full potential of artificial intelligence to be realised whilst mitigating the associated risks.

References:

- Alhudaidi, I., Tomasević, V., Božić, S. 2025. Security implications of AI usage in government strategic decision-making in the UAE. *Security Science Journal*, Vol. 6 No. 2, 138-156.
- Abomakhelb, A., Jalil, K.A., Buja, A.G., Alhammadi, A., Alenezi, A.M. 2025. A Comprehensive Review of Adversarial Attacks and Defence Strategies in Deep Neural Networks. *Technologies*, 13(5).
- Aarab, A., El Marzouki, A., Boubker, O., El Moutaqi, B. 2025. Integrating AI in Public Governance: A Systematic Review. *Digital*, 5(4), 1-36.
- Berko, S. 2025. AI-Driven Threat Detection for Public Sector Systems: Balancing Innovation and Privacy. *World Journal of Innovation and Modern Technology*, Vol. 9, No. 10, 231-254.
- Cheong, B.C. 2024. Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making. *Frontiers in Human Dynamics*, 6, 1-11.
- Ciekanowski, Z., Gruchelski, M., Nowicka, J., Żurawski, S., Pauliuchuk, Y. 2023. Cyberspace as a Source of New Threats to the Security of the European Union. *European Research Studies Journal*, Volume XXVI, Issue 3, 782-797.
- Cybersecurity roles and skills for NIS2 essential and important entities. Mapping NIS2 obligations to ECSF, ENISA 2025.
- Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972 and repealing Directive (EU) 2016/1148 (NIS 2 Directive).
- Grima, S., Thalassinos, E., Cristea, M., Kadłubek, M., Maditinos, D., Peiseniece, L. (Eds.). 2023. *Digital transformation, strategic resilience, cyber security and risk management*. Emerald Publishing Limited.
- He, Z., Davila, D., Bi, S., Wang, T., Hou, T. 2025. Machine Learning for Cybersecurity: A Survey of Applications, Adversarial Challenges, and Future Research Directions. *Electronics*, 14(23).
- Hofmann, B. 2025. Biases in AI: acknowledging and addressing the inevitable ethical issues. *Frontiers Digital Health*, 7, 1-15.
- 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*, Volume 8, Issue 2, 1-15.

- Kaczmarek, K., Karpiuk, M., Spaziani, A. 2024. Use of artificial intelligence in the public sector: threats and prospects. *Studia Iuridica Toruniensia*, vol. XXXVI, 2024, 31–50.
- Okdem, S., Okdem, S., Artificial Intelligence in Cybersecurity: A Review and a Case Study. *Applied Science*, 14, 10487, 2024, 1–20.
- Raksnys, A.V., Gudelis, D., Guogis, A., The Uses of Artificial Intelligence in the Public Sector: Challenges and Prospects. *Public Policy and Administration*, Vol. 24, No. 3, 467-477.
- Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144, and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act).
- Tveita, L.J., Hustad, E. 2025. Benefits and Challenges of Artificial Intelligence in the Public Sector: A Literature Review. *Procedia Computer Science*, Volume 256, 222-229.
- Tyagi, P., Grima, S., Sood, K., Balamurugan, B., Özen, E., Thalassinou, E. (Eds.). 2023. Smart analytics, artificial intelligence and sustainable performance management in a global digitalised economy. Emerald Publishing Limited.
- Zurawski, S., Chrzaszcz, A., Ciekanski, Z., Pauliuchuk, Y., Pietrzyk, S., Wyrzykowska, B. 2025. Effectiveness of Information Security Incident Management Systems: Identifying Practices, Challenges and Development Perspectives. *European Research Studies Journal*, Volume XXVIII, Issue 1, 575-588.