NATIONAL SECURITY
НАЦИОНАЛНА СИГУРНОСТ

AN AI-SUPPORTED REFERENCE PROCESS FOR RISK MANAGEMENT IN ACCORDANCE WITH ISO 27005

Sebastian Grüneberg
University of Library Studies and Information Technologies
https://doi.org/10.70300/YRHH6679

Abstract: *The increasing complexity of information security risks creates significant challenges for organizations in risk management. ISO 27005 provides a normative framework but does not explicitly specify how to implement individual process phases methodologically. While the use of artificial intelligence (AI) in risk management is gaining attention, consistent process-oriented integration has yet to be achieved. This article explores how to systematically integrate AI into the risk management process in accordance with ISO 27005 to ensure compliance with standards. Through document analysis of the standard and a systematic literature review, an AI-supported reference process was developed that positions AI support functions across all process phases and establishes clear decision and validation points. The findings indicate that AI is mainly useful for organizing and analyzing risk-related information, while humans retain responsibility for decisions with normative implications.*
Keywords: *ISO 27005; Artificial Intelligence; Information Security Risk Management; Large-Language-Models (LLM); Reference Process*

INTRODUCTION

The importance of systematic risk management has grown significantly in nearly all organizations in recent years. The ongoing digitalization and interconnectedness of business processes are creating risk environments characterized by high complexity, dynamic dependencies, and increasing amounts of risk-related information. Particularly in information security, traditional manual methods for identifying, analyzing, and prioritizing risks are increasingly reaching their limits. In this context, many scientific studies highlight that conventional risk analysis methods and models are becoming inadequate. These models typically rely on subjective expert assessments and static evaluation frameworks. Such models cannot keep up with the constantly evolving dynamics of modern risk scenarios. Meanwhile, artificial intelligence (AI) has proven to be a powerful tool for analyzing complex data and decision-making situations and is increasingly discussed as a potential aid in risk management.

ISO 27005 offers a well-established framework for information security risk management, but deliberately does not specify particular methods or tools for carrying out individual process steps. This flexibility enables adaptable applications but often leads to diverse approaches and a heavy reliance on personal expertise. Although the number of contributions on the use of AI in risk management is increasing, there is still a lack of a process-oriented, standard-compliant concept that systematically integrates AI into the process phases of ISO 27005 and clearly defines the limits of automation.

This study fills this gap by creating an AI-supported reference process for risk management based on ISO 27005, derived from a systematic literature review. The reference process links AI application areas to specific phases of ISO 27005 and emphasizes key decision points and validation checkpoints. This article intentionally does not aim to maximize automation in risk management. Instead, the reference process functions as a control and governance framework that ensures the standard-compliant, traceable, and responsible integration of AI.

RESEARCH METHODOLOGY

This study adopts a conceptual-analytical approach. Methodologically, it combines a document analysis of the standard with a systematic review of literature on the use of AI in information security and IT risk management. In the first step, ISO 27005 was analyzed to understand the nor-

mative process structure of risk management and identify process phases where the standard intentionally does not specify detailed methodological guidelines. This openness provides the foundation for exploring potential AI integration points within the normative framework. Based on these findings, a systematic literature review was carried out focusing on AI applications in information security and IT risk management. The review included relevant scientific publications that explicitly mention risk analysis, assessment, and management processes. The initial search yielded about 80 relevant papers. Through a multi-stage selection process, non-specialist contributions, purely technical studies lacking process references, and papers not applicable to the context of ISO 27005 were excluded, leaving 47 papers for analysis. During the analysis, the publications were assessed based on the normative process phases of ISO 27005. The individual studies were intentionally not evaluated for quality. Instead, the goal was to create a structured overview of AI use cases described in the literature. The following five analytical perspectives guided the literature review:

(1) Assignment of the contributions to the individual process phases,

(2) Identification of AI-supported tasks and the functional role of AI,

(3) Survey of relevant input and output artefacts,

(4) Identification of explicit decision and validation points, including human-in-the-loop requirements

(5) Recording of the stated prerequisites and limitations of AI use

The results of this analysis were summarized into a cross-phase synthesis, from which the AI-supported reference process was developed. The selected works are chosen to cover a representative range of AI application areas across the ISO 27005 process phases, without trying to be exhaustive regarding individual method classes.

RESULTS

The results of this article include the development of an AI-supported reference process for risk management aligned with ISO 27005. The following sections demonstrate how AI support functions, process artifacts, and decision and validation points can be systematically positioned along the individual process phases. The reference process is thus not presented as a collection of isolated AI use cases but as a coherent, process-oriented framework. It should be viewed as a control and governance structure that defines the conditions, limitations, and validation points for AI use within a normative risk management process. The assignment of AI support functions is based on a detailed assessment of process characteristics. Of particular importance are the normative influence of an activity, its context dependency, the reversibility of the generated results, and the stability and availability of the underlying data.

These differentiation criteria, shown in Table 1, are the result of a cross-phase analysis of the literature and the normative process logic of ISO 27005. They unify recurring arguments about the role of AI, responsibility assignment, and the management of uncertainty and data dependency into a consistent basis for evaluation. Based on this, the reference process distinguishes between AI support and human decision-making.

Table 1. Demarcation criteria AI vs. human

Criterion	Suitable for AI	Human required
Normative influence	Structuring, aggregation	Acceptance, prioritisation
Context dependency	Generic patterns	Organisation-specific meaning
Reversibility	Preliminary results	Binding decisions
Data stability	Structured, historical	Uncertain, incomplete

Figure 1 illustrates the AI-supported risk management process as defined in ISO/IEC 27005. The diagram shows AI support functions at each process stage and highlights decision-making and validation points for human actors.

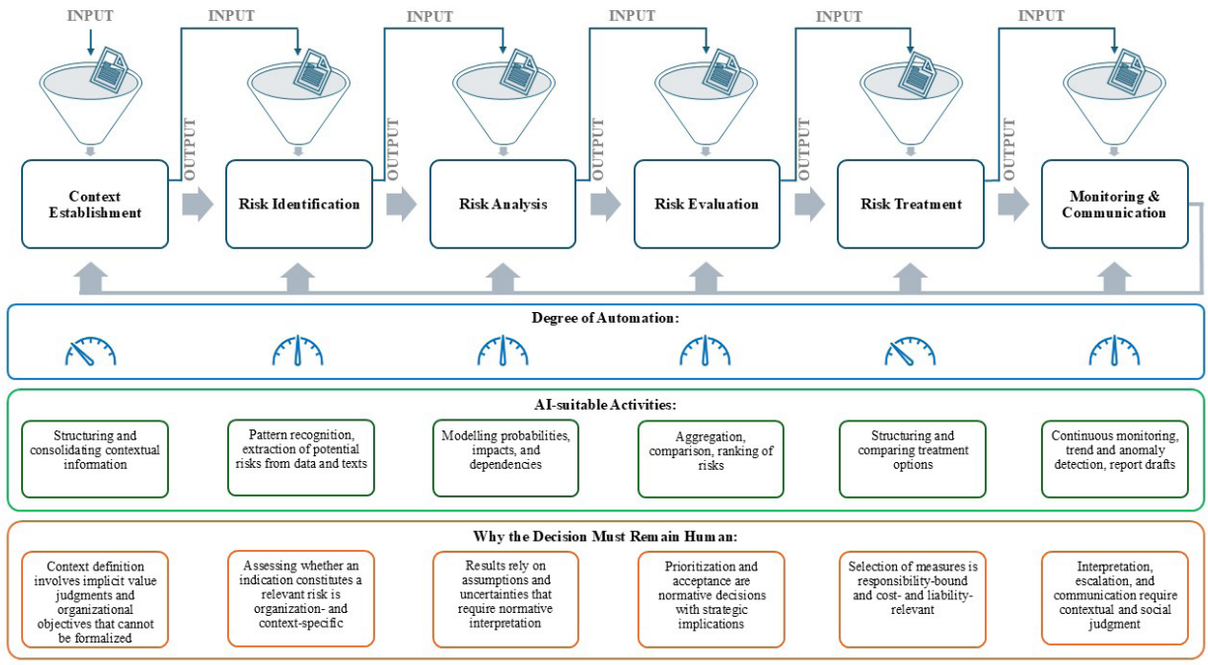

Fig. 1. Reference process (own illustration)

Context establishment forms the starting point of the risk management process according to ISO 27005 and defines the organizational, technical, and regulatory framework within which risks are identified, analyzed, and evaluated. These framework conditions include defining key reference points, such as scopes, risk criteria, and protection requirements, which significantly impact the further course of risk management (DIN Deutsches Institut für Normung e. V., 2024). The examined literature shows that AI methods are primarily used to establish context to support the organization and integration of diverse information. It focuses on describing AI-supported techniques for analyzing large documents, guidelines, system descriptions, and technical artifacts. The goal is to create a clear, understandable starting point for subsequent steps in the process. AI's role is limited to pre-structuring, categorizing, and aggregating activities. In this context, special attention is given to how taxonomies, ontologies, and knowledge graphs are used to ensure consistent mapping of organizational and technical contexts (Lourenço et al. 2025; Szczekocka 2024) and linking governance, risk and control structures (Eisenberg et al. 2025). Additionally, normative and risk theory contributions emphasize that defining context parameters always involves implicit value assumptions that cannot be resolved automatically (Aven 2016). Further work emphasizes the importance of early governance and architecture decisions in defining the context of AI-supported systems and explicitly prioritizes these before addressing the actual risk analysis (Bodnari and Travis 2025; Bogdanov et al. 2024; Tabassi 2023). Kandasamy (2024) also highlights that definitions of context are always shaped by implicit assumptions about system boundaries and knowledge availability. Relevant input artifacts include guidelines, process descriptions, architecture overviews, and asset lists. Structured context information is produced as output artifacts. These serve as a basis for identifying risks. However, defining risk criteria, protection requirements, and organizational scopes requires normative decisions that are directly linked to the organization's strategic objectives and responsibilities. These decisions are highly context-dependent and cannot be fully formalized. In this phase, humans are responsible for the definitive definition of context, while AI only performs preparatory and supportive functions.

In the **risk management** process, according to ISO 27005, risk identification marks the shift from describing the context to conducting operational analysis. During this phase, potential threats, vulnerabilities, and undesirable events are systematically documented to establish an initial foundation for the subsequent risk analysis (DIN Deutsches Institut für Normung e. V., 2019, 2024). The reviewed literature describes the application of AI to risk identification, particularly in

data-heavy settings where large volumes of diverse security-related information must be assessed. AI techniques facilitate automated analysis of internal and external data sources, including logs, monitoring data, vulnerability details, and external threat intelligence. The goal is to generate structured indicators of potential risks by detecting patterns, anomalies, and correlations among indicators. These include, in particular, approaches based on NLP and LLM for extracting risks from unstructured text sources (Begam et al. 2025; Hassani et al. 2024; Liu et al. 2025), as well as knowledge-based and graph-supported techniques for structured risk derivation (Bahr et al. 2025). Additionally, several papers describe using AI as a tool to expand the identification of the risk spectrum, such as through simulation-, evolution-, or scenario-based approaches. These methods are particularly aimed at identifying previously unknown or emerging risks (Chockalingam et al. 2017; Dass and Siami Namin 2020; Kalejaiye 2025; Kioskli et al. 2024; Takko et al. 2023). At the same time, empirical studies show that generative approaches specifically carry risks like hallucinations or misclassifications and therefore do not offer a reliable foundation for autonomous risk assessment (Collier et al. 2025; Esposito et al. 2024). The output artifacts of this process are pre-structured risk indicators or preliminary risk lists used as input for risk analysis. However, these indicators do not represent risks in the normative sense. The decision on whether identified indicators should be regarded as relevant risks depends on their classification within the organizational and regulatory context, as well as on implicit knowledge of business processes and dependencies. This validation cannot be fully automated and remains the responsibility of humans, while AI plays only a decision-preparatory role in risk identification.

As part of the **risk analysis**, the identified risks are evaluated based on their likelihood of occurrence and possible impact. This phase involves systematically organizing and modeling uncertainty, which provides the analytical foundation for the subsequent risk assessment. The examined literature indicates that AI methods are primarily used to support complex analysis and modeling tasks in risk analysis. Data-driven, probabilistic, and model-based approaches are discussed, considering historical incident data, dependency structures, or scenarios. AI consistently processes assessment parameters, organizes relevant influencing factors, and simulates potential developments. Examples include fuzzy-based and Bayesian models for representing uncertainty (Abdymanapov et al. 2021; Savchuk 2023) as well as graph- and deep learning-based methods for modeling risk dependencies and propagation mechanisms (Chhetri and Namin 2025; Shanthi et al. 2025). Additionally, several studies emphasize the importance of explainable models to ensure the traceability of analytical results (Enríquez 2023; Giudici and Raffinetti 2022). Furthermore, other studies employ probabilistic, Bayesian, and hybrid modeling approaches to explicitly address uncertainties, dependencies, and incomplete data sets in risk analysis (Ding and Liu 2012; Lezzi et al. 2025; Wang et al. 2024). The identified risks, relevant contextual information, and historical data serve as input. The output artifacts generated include the calculation of evaluation parameters, the creation of scenarios, and the recording of modeling results. However, interpreting these results requires managing uncertainties, assumptions, and data gaps that cannot be fully formalized. Assessing the significance and relevance of the modeled results relies on organization-specific knowledge and normative beliefs. Therefore, the interpretation and plausibility checking of the analysis results remain the responsibility of humans.

Risk evaluation aims to categorize and rank analysis results based on established criteria, providing a solid foundation for future risk management decisions. It marks the shift from analytical modeling to normative decision-making. The literature reviewed describes how AI is used during this phase, particularly to assist in comparing and structuring risks. AI techniques are employed in various areas, including aggregating assessment outcomes, prioritizing risks, and performing consistency checks. Based on this, AI creates structured summaries or rankings but does not make independent decisions. These approaches range from multi-criteria assessment methods to ensemble-based prioritization models (Campos et al. 2025; Giudici and Raffinetti 2022). Empirical studies show that LLMs, in particular, produce inconsistent results in numerical evaluation and scoring tasks and therefore do not offer a dependable basis for decision-making (Collier et al. 2025). Furthermore, risk theory emphasizes that quantitative risk values must always be interpreted within

the context of the underlying knowledge base (the strength of knowledge). Although AI can model probabilities and impacts, assessing epistemic uncertainty, implicit assumptions, and the robustness of available data requires human judgment (Aven 2016; Erdogan et al. 2021). Additionally, multiple studies have demonstrated that AI-based prioritization heavily relies on the fundamental assessment logic and data quality. Without human plausibility checks, this can lead to systematic biases (Ablayeva 2025; Kunle-Lawanson 2022). Relevant output artifacts include prioritized risk lists and aggregated assessment overviews. However, the decision on risk acceptance, escalation, or treatment remains a human responsibility, as it involves normative, strategic, and organizational aspects that cannot be fully formalized. The role of AI in risk assessment is therefore strictly limited to supporting structured decision-making processes.

Risk treatment involves choosing and applying measures to reduce, transfer, or accept risks. This stage marks the shift from assessment to action and is directly connected to organizational responsibility. The reviewed literature highlights the use of AI in risk treatment mainly to support the analysis and organization of potential measures. AI-supported techniques help, for example, assess measures based on their effectiveness, dependencies, and cost-benefit ratios, and provide an overview of alternative options. These include, in particular, methods for analyzing and optimizing security measures (Eke and Onyejegbu 2024; Finistrella et al. 2025) as well as AI-supported methods for generating and evaluating textual proposals for security measures (Xia et al. 2021). However, several studies emphasize that selecting and implementing measures cannot be automated because of their regulatory, economic, and organizational implications (Collier et al. 2025). Previous work has shown that although algorithmic optimization of policy options can lead to efficiency gains, the final choice must account for organization-specific trade-offs (Kirt and Kivimaab 2010). Relevant output artifacts include structured policy options and comparative overviews. The selection, prioritization, and implementation of specific measures require weighing effectiveness, costs, and feasibility and assessing regulatory requirements. This weighing is highly context-dependent and involves organizational responsibility and potential consequences. Accordingly, these decisions remain the responsibility of humans, with AI playing only a supporting role.

Risk monitoring and communication ensure that risks are continuously observed, assessed, and addressed to relevant stakeholders. This phase does not follow a sequential process but serves as ongoing feedback for all previous steps. The reviewed literature describes the use of AI in this phase, notably to support the continuous evaluation and processing of risk-related information. AI techniques can analyze real-time monitoring data, update risk assessments, and automatically generate reports or alerts. Additionally, learning-based approaches to anomaly and attack detection are discussed (Joshi and Shandilya 2024; Rihan et al. 2023; Yang et al. 2023) as well as explainable AI models to assist in alarm interpretation (Dilhara 2025; Zhang et al. 2022). Multiple studies show that without human validation, false alarms and alarm fatigue can happen (Almadhor et al. 2025; Sharma et al. 2025). Studies on the use of AI-based monitoring and alerting systems show that, especially in complex system landscapes, human interpretation remains necessary to prevent false alarms and contextless escalations (Alexandrov 2025; Korniszuk and Sawicki 2024). Humans remain responsible for assessing consequences, adapting measures, and communicating risk-relevant decisions. Particularly in communication, the literature highlights the importance of transparency, traceability, and interpretation suitable for the target audience. These requirements extend beyond the analytical capabilities of AI systems, making human control and responsibility essential in this phase.

The cross-phase evaluation demonstrates that the use of AI in risk management, as defined in ISO 27005, can be systematically differentiated along functional and normative boundaries. While AI helps structure, analyze, and consolidate risk-relevant information in all process phases, decisions with normative implications remain the responsibility of humans throughout. The reference process derived from this clearly shows this distinction by consistently placing AI support functions, process artifacts, and decision and validation points across all process phases. The reference process is therefore not meant as an architecture for automating risk decisions but as a framework for structuring and controlling the norm-compliant integration of AI into risk management, in accordance with ISO 27005.

DISCUSSION

The results of the study show that AI can be effectively used in risk management, in line with ISO 27005, as part of a process-oriented approach, provided its role is clearly defined and integrated into existing governance structures. Organizing the literature findings by individual process phases makes it clear that AI should not be seen as an isolated technology but rather as a supporting element within a normative risk management process. At the same time, the review of the existing literature reveals significant variability in the attribution of automation levels and responsibilities. This variability is less due to fundamental deficits in AI processes and more to differences in data situations, application goals, and organizational contexts. The developed reference process addresses these findings by consistently positioning AI as a decision-preparatory element and explicitly including human decision-making and validation points in all process phases.

Regarding LLMs in particular, current research increasingly highlights specific limitations. Empirical studies indicate that when LLMs are used in numerical assessments and scoring tasks, such as in FMEA-like procedures, results are often inconsistent or methodologically unreliable (Collier et al. 2025). While such models can efficiently organize textual information and describe risks linguistically, their effectiveness for quantitative risk assessments is still limited. Meanwhile, studies show that human experts generally offer more accurate and context-aware evaluations. AI systems, however, excel in speed, scalability, and providing action-oriented preliminary structuring, such as aggregated, prioritized, or similarly prepared information. (Esposito et al. 2024). These findings support the distinction made in the reference process between AI as an assistant and not as a decision-making authority. Furthermore, the results show that the use of AI is always linked to organizational, normative, and governance-related framework conditions. The reference process intentionally does not prescribe a single set of methods but allows for a context-dependent selection of suitable AI procedures. This flexibility reflects the different requirements of individual process phases and varying levels of maturity and data availability across organizations. Methodological diversity is viewed not as a weakness but as a necessary condition for the practical integration of AI. Another important aspect concerns the traceability and explainability of AI-supported results. Non-transparent models pose the risk that results will either be accepted uncritically or rejected outright. The reference process addresses this tension by positioning explainability not as an extra technical feature but as a procedural necessity for risk monitoring, communication, and governance. Consequently, explainability becomes essential to responsible decision support.

The article adds conceptual value by turning existing, often fragmented AI application areas into a consistent, standardized process logic. However, the reference process has limitations: it relies on a synthesis of the literature and has not been empirically tested in specific organizational settings. Different levels of maturity, data availability, and regulatory environments necessitate organization-specific adjustments. Promising directions for future research include empirical evaluations of the reference process, detailed analyses of individual process phases, and studies of specific regulatory or cultural conditions. Alternative approaches that aim for extensive automation of risk assessments or decisions are critically evaluated in light of the results of this article. These approaches often implicitly embed normative assumptions into models without making them transparent or verifiable. Under ISO 27005, however, risk decisions are closely tied to organizational responsibility and accountability. Therefore, the developed reference process consistently positions AI as a supporting, rather than a primary, element within normatively guided risk management.

CONCLUSIONS

This article explores the possibility of systematically integrating artificial intelligence into the risk management process in line with ISO 27005 without compromising normative principles, responsibilities, and decision-making structures. Based on a document review of the standard and a systematic review of the literature, an AI-supported reference process was developed that consistently assigns AI support functions to the different process phases while also establishing clear boundaries for automation.

The study's findings indicate that applying AI in risk management is especially important when a large amount of diverse information needs to be processed into a clear, analytical, and concise format. However, decisions with normative implications, particularly regarding the assessment, acceptance, and handling of risks, must still be made by humans. The reference process explicitly enforces this distinction and makes the decision and validation points visible throughout the entire process phases.

The main scientific contribution is not evaluating individual AI methods but rather the process-oriented synthesis of existing research into a norm-based control-and-governance framework. Therefore, the reference process should be seen not as an architecture for automated decision-making but as a guideline for the structured, transparent, and responsible integration of AI into risk management. Since the reference process is built on a synthesis of the literature, empirical studies on its practical application and context-specific adaptations in different organizational and regulatory environments are natural starting points for future research.

REFERENCES

ABDYMANAPOV, S. A., MURATBEKOV, M., ALTYNBEK, S. and BARLYBAYEV, A., 2021. Fuzzy Expert System of Information Security Risk Assessment on the Example of Analysis Learning Management Systems. *IEEE Access*, vol. 9, pp. 156556–156565.

ABLAYEVA, O., 2025. Comparative analysis of random forest, SVM, and LSTM algorithms for threat detection in internet domains. *International Journal of Artificial Intelligence*, vol. 5, no. 5, pp. 1499–1504. [online]. Available from: https://www.academicpublishers.org/journals/index.php/ijai/article/view/4679

ALEXANDROV, A., 2025. LSTM-RNN Method for Anomaly-Based Intrusion Detection Systems. [online]. Available from: https://ceur-ws.org/Vol-3971/paper02.pdf [viewed 29 December 2025].

ALMADHOR, A., ALSUBAI, S., KRYVINSKA, N., AL HEJAILI, A., BOUALLEGUE, B., AYARI, M. and ABBAS, S., 2025. Transfer Learning with XAI for Robust Malware and IoT Network Security. *Scientific Reports*, vol. 15, no. 1, p. 26971.

AVEN, T., 2016. Risk Assessment and Risk Management: Review of Recent Advances on Their Foundation. *European Journal of Operational Research*, vol. 253, no. 1, pp. 1–13.

BAHR, L., WEHNER, C., WEWERKA, J., BITTENCOURT, J., SCHMID, U. and DAUB, R., 2025. Knowledge Graph Enhanced Retrieval-Augmented Generation for Failure Mode and Effects Analysis. *Journal of Industrial Information Integration*, vol. 45, p. 100807.

BEGAM, F., MANIMEKALAI, K. and RANGASAMY, B., 2025. Transformer Models and Attention Mechanisms for Intelligent Cyber Threat Intelligence Extraction, pp. 325–354.

BODNARI, A. and TRAVIS, J., 2025. Scaling Enterprise AI in Healthcare: The Role of Governance in Risk Mitigation Frameworks. *NPJ Digital Medicine*, vol. 8, no. 1, p. 272.

BOGDANOV, D., ETTI, P., KAMM, L. and STOMAKHIN, F., 2024. Artificial Intelligence System Risk Management Methodology Based on Generalized Blueprints, pp. 123–140.

CAMPOS, S., PAPADATOS, H., ROGER, F., TOUZET, C., QUARKS, O. and MURRAY, M., 2025. A Frontier AI Risk Management Framework: Bridging the Gap Between Current AI Practices and Established Risk Management.

CHHETRI, B. and NAMIN, A. S., 2025. The Application of Transformer-Based Models for Predicting Consequences of Cyber Attacks.

CHOCKALINGAM, S., PIETERS, W., TEIXEIRA, A. and VAN GELDER, P., 2017. Bayesian Network Models in Cyber Security: A Systematic Review. In: LIPMAA, H., MITROKOTSA, A. and MATULEVIČIUS, R., eds. *Secure IT Systems*. Cham: Springer International Publishing, pp. 105–122.

COLLIER, Z. A., GRUSS, R. J. and ABRAHAMS, A. S., 2025. How Good Are Large Language Models at Product Risk Assessment? *Risk Analysis*, vol. 45, no. 4, pp. 766–789.

DASS, S. and SIAMI NAMIN, A., 2020. Evolutionary Algorithms for Vulnerability Coverage. In: *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*. Madrid, Spain, 13–17 July 2020. Piscataway, NJ: IEEE, pp. 1795–1801.

DILHARA, A., 2025. Impact of Machine Learning and AI on Cybersecurity Risks and Opportunities.

DIN DEUTSCHES INSTITUT FÜR NORMUNG E. V., 2019. *DIN EN IEC 31010:2019 Risk Management – Risk Assessment Techniques*. Berlin: Beuth Verlag.

DIN DEUTSCHES INSTITUT FÜR NORMUNG E. V., 2024. *DIN ISO/IEC 27005:2024 Information Security, Cybersecurity and Privacy Protection – Guidance on Managing Information Security Risks*. Berlin: Beuth Verlag.

DING, D. and LIU, X., 2012. Bayesian Methods with Application in Risk Analysis.

EISENBERG, I. W., GAMBOA, L. and SHERMAN, E., 2025. The Unified Control Framework: Establishing a Common Foundation for Enterprise AI Governance, Risk Management and Regulatory Compliance.

EKE, D. R. N. and ONYEJEGBU, P. L. N., 2024. A Deep Reinforcement Learning Deep Q Network-Based Approach

for Systematic Detection of Distributed Denial of Service Attacks in Software-Defined Networking.
ENRÍQUEZ, L., 2023. A Quantitative Approach to Artificial Intelligence Legal Risk Management. In: SPINDLER, G. and MURIEL CICERI, J. H., eds. *Challenges of Law and Technology – Herausforderungen des Rechts und der Technologie – Retos del Derecho y de la Tecnología*. Göttingen: Göttingen University Press, pp. 169–182.
ERDOGAN, G., GARCIA-CEJA, E., HUGO, Å., NGUYEN, P. H. and SEN, S., 2021. A Systematic Mapping Study on Approaches for AI-Supported Security Risk Assessment. In: *2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC)*. Piscataway, NJ: IEEE, pp. 755–760.
ESPOSITO, M., PALAGIANO, F., LENARDUZZI, V. and TAIBI, D., 2024. Beyond Words: On Large Language Models Actionability in Mission-Critical Risk Analysis. In: *Proceedings of the 18th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*. Barcelona, Spain, 24–25 October 2024. New York, NY: ACM, pp. 517–527.
FINISTRELLA, S., MARIANI, S. and ZAMBONELLI, F., 2025. Multi-Agent Reinforcement Learning for Cybersecurity: Classification and Survey. *Intelligent Systems with Applications*, vol. 26, p. 200495.
GIUDICI, P. and RAFFINETTI, E., 2022. Explainable AI Methods in Cyber Risk Management. *Quality and Reliability Engineering International*, vol. 38, no. 3, pp. 1318–1326.
HASSANI, I. E., MASROUR, T., KOUROUMA, N., MOTTE, D. and TAVČAR, J., 2024. Integrating Large Language Models for Improved Failure Mode and Effects Analysis (FMEA): A Framework and Case Study. *Proceedings of the Design Society*, vol. 4, pp. 2019–2028.
JOSHI, I. P. and SHANDILYA, V. K., 2024. A Critical Review – Use of Ensemble Methods in Intrusion Detection System. *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 21S, pp. 1157–1164. [online]. Available from: https://ijisae.org/index.php/IJISAE/article/view/5517
KALEJAIYE, A. N., 2025. Federated Learning in Cybersecurity: Privacy-Preserving Collaborative Models for Threat Intelligence Across Geopolitically Sensitive Organizational Boundaries. *International Journal of Research Publication and Reviews*, vol. 6, no. 7, pp. 227–249.
KANDASAMY, U. C., 2024. Ethical Leadership in the Age of AI: Challenges, Opportunities and Framework for Ethical Leadership.
KIOSKLI, K., RAMFOS, A., TAYLOR, S., MAGLARAS, L. and LUGO, R., 2024. Optimizing AI System Security: An Ecosystem Recommendation to Socio-Technical Risk Management. In: *Human Factors in Design, Engineering, and Computing*. AHFE International.
KIRT, T. and KIVIMAAB, J., 2010. Optimizing IT Security Costs by Evolutionary Algorithms. In: *Conference on Cyber Conflict Proceedings*, pp. 145–160. [online]. Available from: https://ccdcoe.org/uploads/2018/10/1_Proceedings2010FullBook.pdf [viewed 29 December 2025].
KORNISZUK, K. and SAWICKI, B., 2024. Autoencoder-Based Anomaly Detection in Network Traffic, pp. 1–4.
KUNLE-LAWANSON, O., 2022. The Role of AI in Information Security Risk Management. *World Journal of Advanced Engineering Technology and Sciences*, vol. 7, no. 2, pp. 308–319.
LEZZI, M., MONTEFUSCO, P., LAZOI, M. and CORALLO, A., 2025. AI-Based Cybersecurity for a Sustainable Digital Industry: Systematic Literature Review and Future Research Directions. *Journal of Industrial Information Integration*, vol. 48, p. 100980.
LIU, G., LU, K. and PI, S., 2025. Graph Neural Networks Embedded with Domain Knowledge for Cyber Threat Intelligence Entity and Relationship Mining. *PeerJ Computer Science*, vol. 11, e2769.
LOURENÇO, B., ADÃO, P., FERREIRA, J. F., MARQUES, M. M. and VAZ, C., 2025. Structuring Security: A Survey of Cybersecurity Ontologies, Semantic Log Processing, and LLM Applications.
RIHAN, S. D. A., ANBAR, M. and ALABSI, B. A., 2023. Meta-Learner-Based Approach for Detecting Attacks on Internet of Things Networks. *Sensors*, vol. 23, no. 19.
SAVCHUK, V., 2023. Bayesian Risk Assessment Technique for Economic Stress-Strength Models. [online]. Available from: https://mpra.ub.uni-muenchen.de/119078/ [viewed 29 December 2025].
SHANTHI, S., S, S. S. and SINDHU, M., 2025. Graph-Based Risk Analysis Using Graph Neural Networks for Mapping Cyber Threat Propagation in Large-Scale Networks, pp. 72–106.
SHARMA, A., RANI, S. and SHABAZ, M., 2025. A Comprehensive Review of Explainable AI in Cybersecurity: Decoding the Black Box. *ICT Express*, vol. 11, no. 6, pp. 1200–1219.
SZCZEKOCKA, E., 2024. AI Risk Management Implementation Challenges. *Computer Science & Information*, vol. 14, no. 17. [online]. Available from: https://aircconline.com/csit/papers/vol14/csit141710.pdf [viewed 29 December 2025].
TABASSI, E., 2023. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. Gaithersburg, MD: National Institute of Standards and Technology (NIST).
TAKKO, T., BHATTACHARYA, K., LEHTO, M., JALASVIRTA, P., CEDERBERG, A. and KASKI, K., 2023. Knowledge Mining of Unstructured Information: Application to Cyber Domain. *Scientific Reports*, vol. 13, no. 1, p. 1714.
WANG, L., CHENG, Y., XIANG, A., ZHANG, J. and YANG, H., 2024. Application of Natural Language Processing in Financial Risk Detection.
XIA, Y., JAZDI, N. and WEYRICH, M., 2021. Enhance FMEA with Large Language Models for Assisted Risk Man-

agement in Technical Processes and Products. *Computers in Industry*, vol. 131.
YANG, A., LU, C., LI, J., HUANG, X., JI, T., LI, X. and SHENG, Y., 2023. Application of Meta-Learning in Cyberspace Security: A Survey. *Digital Communications and Networks*, vol. 9, no. 1, pp. 67–78.
ZHANG, Z., HAMADI, H. A., DAMIANI, E., YEUN, C. Y. and TAHER, F., 2022. Explainable Artificial Intelligence Applications in Cyber Security: State-of-the-Art in Research. *IEEE Access*, vol. 10, pp. 93104–93139.

ПОДПОМАГАН ОТ ИЗКУСТВЕН ИНТЕЛЕКТ РЕФЕРЕНТЕН ПРОЦЕС ЗА УПРАВЛЕНИЕ НА РИСКА В СЪОТВЕТСТВИЕ С ISO 27005

Резюме: *Нарастващата сложност на рисковете за информационната сигурност създава значителни предизвикателства за организациите в управлението на риска. ISO 27005 предоставя нормативна рамка, но не посочва изрично как да се прилагат методологично отделните фази на процеса. Въпреки че използването на изкуствен интелект (ИИ) в управлението на риска привлича все по-голямо внимание, все още не е постигната последователна интеграция, ориентирана към процесите. В тази статия се разглежда как систематично да се интегрира ИИ в процеса на управление на риска в съответствие с ISO 27005, за да се гарантира съответствие със стандартите. Чрез анализ на документацията на стандарта и систематичен преглед на литературата е разработен референтен процес, подкрепен от ИИ, който позиционира ИИ поддържащите функции във всички фази на процеса и установява ясни точки за вземане на решения и валидиране. Резултатите показват, че ИИ е полезен главно за организиране и анализиране на информация, свързана с риска, докато хората запазват отговорността за решенията с нормативни последствия.*
Ключови думи: *ISO 27005; изкуствен интелект; управление на риска за информационната сигурност; големи езикови модели (LLM); референтен процес*

Себастиан Грюнеберг, докторант
Researcher ID: 0009-0002-2339-0729 (ORCID)
Университет по библиотекознание и информационни технологии
София, България
E-mail: Sebastian.Grueneberg@outlook.com