applied sciences

MDPI

*Article*

# Development of a Flexible Information Security Risk Model Using Machine Learning Methods and Ontologies

**Alibek Barlybayev [1,2,*], Altynbek Sharipbay [2], Gulmira Shakhmetova [2] and Ainur Zhumadillayeva [2]**

[1] Higher School of Information Technology and Engineering, Astana International University, Astana Z05H0T3, Kazakhstan
[2] Research Institute of Artificial Intelligence, L.N. Gumilyov Eurasian National University, Astana Z00T8E0, Kazakhstan; sharalt@mail.ru (A.S.); sh_mira2004@mail.ru (G.S.); ay8222@mail.ru (A.Z.)
* Correspondence: frank-ab@mail.ru

**Abstract:** This paper presents a significant advancement in information security risk assessment by introducing a flexible and comprehensive model. The research integrates established standards, expert knowledge, machine learning, and ontological modeling to create a multifaceted approach for understanding and managing information security risks. The combination of standards and expert insights forms a robust foundation, ensuring a holistic grasp of the intricate risk landscape. The use of cluster analysis, specifically applying k-means on information security standards, expands the data-driven approach, uncovering patterns not discernible through traditional methods. The integration of machine learning algorithms in the creation of information security risk dendrogram demonstrates effective computational techniques for enhanced risk discovery. The introduction of a heat map as a visualization tool adds innovation, facilitating an intuitive understanding of risk interconnections and prioritization for decision makers. Additionally, a thesaurus optimizes risk descriptions, ensuring comprehensiveness and relevance despite evolving terminologies in the dynamic field of information security. The development of an ontological model for structured risk classification is a significant stride forward, offering an effective means of categorizing information security risks based on ontological relationships. These collective innovations enhance understanding and management of information security risks, paving the way for more effective approaches in the ever-evolving technological landscape.

**Keywords:** information security risk; assessment model; information security standards; k-means; dendrogram; heat map; thesaurus; ontological model



**Citation:** Barlybayev, A.; Sharipbay, A.; Shakhmetova, G.; Zhumadillayeva, A. Development of a Flexible Information Security Risk Model Using Machine Learning Methods and Ontologies. *Appl. Sci.* **2024**, *14*, 9858. https://doi.org/10.3390/app14219858

Academic Editor: Jose María Alvarez Rodríguez

Received: 26 September 2024
Revised: 17 October 2024
Accepted: 25 October 2024
Published: 28 October 2024



## 1. Introduction

In today's data-driven society and economy, the importance of cyber and information security cannot be overstated. As managers strive to deliver value-based services to their clients, the need for safeguarding critical assets becomes paramount [1]. Unfortunately, many organizations are not adequately investing in protecting valuable information. The fast-paced progress of information and communications technologies is leading to continuous advancements in information systems. However, this also raises concerns about increased vulnerability [2]. Addressing these challenges becomes increasingly complex as the system becomes more intricate with interdependent components [3]. The COVID-19 pandemic has accelerated the growth of digital technologies, placing additional pressure on organizations with inadequate security policies for their information technology (IT)systems. Simultaneously, there are several risk management frameworks available that organizations can utilize to evaluate and effectively handle information security (IS) risks. These frameworks are designed to support organizations in assessing and mitigating potential risks related to their information security [4].

In order to keep up with the ever-changing technological landscape, it is crucial for companies to adapt and embrace transformational shifts. However, they often face a

critical challenge when it comes to existing IT security standards, as they may not be dynamically flexible enough to accommodate new technologies [5]. The current standards not only fall short in addressing the intricate nature of advanced technology but their implementation also comes with significant costs and reduced effectiveness in managing IT risks [6]. With the expansion of digitalization, organizations are confronted with an escalating number of information technology threats. This situation has been further aggravated in the post-COVID-19 landscape [7–9]. The evident shortcomings of current IT risk management frameworks are becoming more apparent. There are several noteworthy drawbacks associated with this approach. These include the use of outdated methodologies, complex implementation processes, an excessive emphasis on compliance, and the need for hybrid approaches to address gaps between controls and compliance requirements [10,11]. Given the circumstances, it is essential to conduct a comprehensive evaluation of the efficacy of existing IT risk management frameworks. Despite extensive research on IT risk management, there remains a notable gap in the literature. Many studies have concentrated on specific contexts, leaving a lack of comprehensive analysis regarding the effectiveness of current IT risk management frameworks [12–14]. The absence of a systematic analysis presents an intriguing research challenge, which this paper aims to address.

The contributions of paper are as follows:

- The research successfully integrates machine learning methods, ontological modeling, and expert knowledge with established standards, creating a multifaceted approach to understanding and managing information security risks. This holistic approach allows for a more robust foundation to address the complex and evolving risk landscape.
- The application of k-means clustering to analyze information security standards and the use of a heat map for visualizing risk interconnections are notable advancements. These methods uncover patterns and relationships not discernible through traditional risk assessment techniques, enhancing the predictive capabilities of the risk model.
- The development of a comprehensive thesaurus and an ontological model for information security risks significantly improves the precision and adaptability of risk descriptions. This innovation ensures the model remains relevant and effective amidst the rapidly changing terminologies and dynamics within the field.
- The model demonstrates exceptional flexibility and adaptability, crucial for contemporary cybersecurity needs. It supports dynamic updates to incorporate new risks and adapts to changing threat landscapes, making it a valuable tool for organizations aiming to enhance their information security posture.

The Introduction section discusses the significance of information security in today's digitalized world and introduces the need for a flexible risk management model. It outlines the paper's goal to develop a new model that integrates various methodologies for a comprehensive risk management approach. The Surveys on IS Threats sub-section reviews the existing literature on information security threats and the importance of robust IT risk management frameworks. The Surveys on IS Risk Management Frameworks sub-section compares and contrasts different frameworks, focusing on those developed by NIST and ISO, to highlight their features and limitations. The Materials and Methods section details the process of examining various information security standards to establish a foundation for the model. The Results section describes the methods used for gathering and preparing data for analysis, emphasizing the importance of a structured approach, utilizes -k-means clustering to identify patterns in information security standards and employs heat maps for the visual representation of data interrelationships; discusses the use of dendrograms to explore data hierarchies and the development of a thesaurus to standardize risk descriptions; presents the outcomes of the cluster analysis and the application of machine learning techniques to categorize information security risks; and evaluates the effectiveness of visual tools like heat maps and dendrograms in illustrating and prioritizing security risks. The Discussion section critically analyzes the strengths and weaknesses of the developed model, comparing it with existing frameworks, and discusses the practical applications of the model in various industries and its potential impact on

enhancing information security risk management. The Conclusions section recaps the major discoveries and innovations introduced in the paper and proposes areas for further investigation and potential improvements to the model.

### 1.1. Surveys on IS Threats

The significance of strong IT risk management frameworks for companies is often emphasized in our modern, digitalized world where everything is interconnected. These frameworks play a crucial role in identifying, evaluating, addressing, and overseeing risks within information systems [15]. By adopting a well-established framework, organizations can gain numerous advantages. One such benefit is the ability to create a prioritized roadmap towards enhancing IT security practices. Additionally, utilizing a common language provided by the framework facilitates effective communication and discussion of IT risk challenges, establishment of security standards for forthcoming legal judgments, as well as the encouragement of proactive IT risk management instead of reactive compliance, is crucial [16]. Nevertheless, the fundamental principles of security and risk serve as the foundation for all IT security frameworks. The protection and safeguarding of information systems are commonly evaluated based on their ability to maintain confidentiality, integrity, and availability, as outlined in [17]. The realm of IT security threats is extensive and continuously changing. These threats encompass a wide range of activities, such as reconnaissance and gathering information [18], phishing attacks, the creation of spoof websites [19], the production of counterfeit certificates [20], and the delivery of malware to internal information systems [21]. The presence of such threats presents substantial obstacles to upholding the security of information systems.

Information security pertains to the safeguarding of information by implementing risk management processes to ensure its confidentiality, integrity, and availability. It aims to protect against unauthorized access, disclosure, alteration, destruction, and disruption of information [22]. Risk management encompasses the crucial process of recognizing, evaluating, and regulating potential vulnerabilities that could jeopardize digital assets such as information, networks, and systems [1]. Risk assessment is a crucial component of effective risk management. It entails the identification of potential hazards, vulnerabilities, and threat vectors. Subsequently, the impact and likelihood of these identified risks are evaluated. The purpose of this process is to establish a foundation for making informed decisions regarding risk mitigation strategies [23]. Although there are various risk management frameworks available, their primary goal remains to safeguard information assets by minimizing risk to an acceptable level while maximizing the business value of the organization [5].

Adversarial threats refer to the risks that arise from individuals or organizations who aim to exploit the dependence on information systems and resources [24]. These individuals can be categorized as external or internal, depending on their role or position. Organizations face various types of threats, including accidental ones that arise from mistakes made by users or administrators. These threats can come from competitive organizations, suppliers, partners, or even customers themselves. It is crucial to take proactive measures to mitigate these risks and safeguard the integrity of your business operations. Structural threats refer to equipment and control failures that occur unexpectedly due to circumstances beyond the usual operating parameters [25]. These failures can be attributed to factors such as resource depletion, equipment aging, or software malfunction. It is crucial to address these threats proactively in order to maintain the reliability and efficiency of your systems. Structural threats encompass a wide range of critical components in any system, including storage, processing, communications, display, sensors, controllers, power supply, operating systems, networking infrastructure, and specialized software. It is crucial to address these threats systematically to ensure the stability and security of your infrastructure. Environmental threats encompass a range of external factors that can have a major impact on an organization's operations [26]. These include natural disasters and infrastructure failures, which are crucial to the smooth functioning of the organization. Recognizing

and mitigating these threats is essential for ensuring business continuity and resilience. Potential disruptions to critical infrastructure can take various forms, ranging from technical glitches like telecommunication failures and power outages to more severe situations such as natural disasters or deliberate acts of destruction. These incidents, including fires, floods, hurricanes, earthquakes, and bombings, have the potential to disrupt essential services and operations.

### 1.2. Surveys on IS Risk Management Frameworks

In today's digital age, the close connection between information systems and organizational processes calls for a comprehensive risk management framework. By integrating these two elements, businesses can effectively mitigate risks and safeguard their operations. It is essential to have a well-defined framework in place to ensure the smooth functioning of both information systems and organizational processes while minimizing potential threats. There are numerous frameworks available in the market, each with its own unique features. However, two of the most widely adopted frameworks used by organizations are those developed by NIST and ISO. The National Institute of Standards and Technology has made significant contributions to the establishment of IT security standards. Two notable standards are NIST SP 800-39 and NIST SP 800-30 (Revision 1). NIST SP 800-39 offers a thorough framework for managing information security risks [27]. The primary objective of the standard is to effectively manage risk at the organizational level. It provides a consistent approach that fosters better governance and enables a deeper comprehension of how IT security influences organizational operations [28].

The standard known as SP 800-39 outlines risk management using four key components: risk framing, risk assessment, risk response strategy, and risk monitoring [29]. The goal of risk framing is to establish an actionable strategy for effectively managing IT security risks. This particular step takes into account the element of risk within the established environment, which acts as the framework for making decisions based on risk factors [30]. Risk framing entails the identification of risks based on various factors such as assumptions, constraints, tolerance levels, priorities, and trade-offs. Specifically, risk tolerance refers to the level of risk that is considered acceptable [27]. Risk framing refers to the process of developing a risk management strategy encompassing risk assessment, risk monitoring, and response strategies. During the risk assessment phase, potential threats, vulnerabilities, possible damages, and likelihood of exploits are identified. The SP 800-30 (Revision 1) standard is the NIST framework that corresponds to IT security risk assessment. In addition, a risk management strategy includes risk monitoring, which is responsible for ensuring consistent compliance verification and assessing the effectiveness of ongoing risk responses [31]. The risk assessment component is a crucial element of risk management as it accurately evaluates the likelihood and potential impact of an IT security risk [32]. The mentioned standard encompasses guidance on the necessary steps to effectively prepare for and carry out a risk assessment. It also includes information on how to monitor and evaluate the assessment processes. It is worth noting that risk assessment is closely connected to the three tiers of risk management as outlined in the SP 800-39 standard [30].

The International Organization for Standardization (ISO) and the International Electro Technical Commission (IEC) have put forth standards to address IT security, known as the ISO/IEC 27000 series. Of particular importance is the ISO/IEC 27005 standard, which provides guidance on information risk security. A proposed approach suggests a continuous process of activity sequences that encompasses several essential steps. These steps involve establishing context, assessing information, treating and monitoring risks, and informing the stakeholders of the organization [29]. The ISO/IEC 27005 standard offers broad guidelines for IT risk management, making it relevant and applicable to a range of organizations, such as small and medium enterprises, nonprofit organizations, and government agencies [33]. One notable distinction between it and the NIST standard lies in its structural framework, specifically in the absence of any prescribed approach for

risk management. Alongside ISO/IEC 27005, the ISO/IEC framework offers guidelines and protocols for adopting a comprehensive approach to implementing, monitoring, and enhancing IT security that aligns with overall organizational risk management practices.

To begin with, it is essential to establish the risk management context. This will help in defining the criteria for identifying risks, allocating responsibility, assessing potential consequences, and assessing the accessibility of information, as well as to establish a systematic approach for evaluating the potential risks and their probabilities [34]. Risk assessment, the second step in the process, involves various stages that include risk identification, risk analysis, and risk evaluation [35]. The framework's iterative nature signifies that the risk assessment process is repeated whenever the information obtained from a previous risk assessment does not adequately support decision making in risk management [36]. An additional cycle of the risk assessment phase is performed, involving updated criteria for evaluating risks or their potential impact [35].

The outcome of a risk assessment process is the development of a risk treatment plan. This plan aims to address identified risks through various strategies, including avoiding, modifying, sharing, or retaining the risk [37]. The risk treatment process is a cyclical one that encompasses multiple procedures. These procedures include evaluating the treatment, estimating the remaining risk levels, modifying the treatment if these levels are deemed unacceptable, and assessing the effectiveness of the treatment according to ISO guidelines. This enables the modification of contextual parameters, including risk acceptance criteria, and the creation of a new version of the treatment. According to the ISO 27005 framework, effective communication and consultation should be maintained throughout all stages of the risk assessment process [38]. Specifically, it is important to effectively communicate risks and treatments to operational staff and managers in order to minimize risks and mitigate potential damage. The framework emphasizes that the controls should be based on risk assessment. This statement effectively reflects the ever-changing nature of risks and emphasizes the significance of the risk monitoring stage. Additionally, the standard furnishes details regarding common threats, organizational constraints, scope considerations, asset valuation, vulnerability assessment, and risk modification.

The Committee of Sponsoring Organizations of the Treadway Commission (COSO) framework is geared towards a broader view of risk management. It integrates IS risks into enterprise risk management, highlighting how IS risks align with organizational objectives and governance.

Control Objectives for Information and Related Technologies (COBIT) is a framework designed for IT governance and management. It provides an actionable set of controls over information technology and aligns IT activities with business objectives through its comprehensive set of processes and best practices.

Comparative analysis of each framework offers unique strengths and caters to different aspects of IS risk management. For example, while NIST provides a security-centric, customizable approach, ISO/IEC 27005 focuses on a structured, compliance-oriented risk management process. COBIT, on the other hand, bridges the gap between IT operational processes and organizational strategic objectives, making it integral for governance and business alignment. A detailed comparative analysis reveals the situational advantages of each framework, depending on organizational size, sector-specific requirements, and regulatory compliance needs.

Tables 1 and 2 summarize the major IS risk management frameworks. These frameworks are crucial for effective risk management and safeguarding operations in organizations across various industries.

One potential alternative approach to establishing security requirements is to integrate them into the requirements engineering process as an integral part of project development. This allows for a more comprehensive and systematic consideration of security needs right from the early stages of the project. By considering security requirements in conjunction with other project requirements, potential vulnerabilities can be identified and addressed proactively, leading to a more secure final product. The Security Quality Requirements

Engineering (SQUARE) methodology was created at Carnegie Mellon University as a framework for effectively identifying and prioritizing the security requirements of IT systems [39]. One significant benefit of this approach is that security concepts are integrated into the initial stages of system development [1]. The SQUARE model consists of nine essential steps. These include identifying safety and security objectives, categorizing requirements based on their level, conducting risk assessments, prioritizing and inspecting requirements [40].

The main results of this [1] scientific paper include the development and evaluation of a fuzzy expert system for information security risk assessment in learning management systems. The authors demonstrate that their approach outperforms traditional methods in terms of accuracy and effectiveness. By utilizing fuzzy logic, the system effectively handles uncertainty and imprecision inherent in risk assessment processes. The novelty of this research lies in the application of a fuzzy expert system specifically tailored for analyzing information security risks in learning management systems. While existing approaches often rely on quantitative or qualitative methods, this paper introduces a novel methodology that leverages fuzzy logic to capture and represent imprecise knowledge related to security risks.

The authors of this paper [41] develop a security assessment framework for CPDPS under intrusion attacks. This framework considers the impact of different types of attacks on the power system's security, reliability, and stability. The novelty of this scientific paper lies in its thorough analysis of the security assessment of cyber–physical distribution power systems under intrusion attacks. The authors provide a comprehensive framework for evaluating the security of such systems and introduce several security metrics to help identify vulnerabilities and improve the overall security of the power grid.

**Table 1.** Traditional IS risk management frameworks.

| Framework | Description | Key Components |
| --- | --- | --- |
| NIST SP 800-39 | Focuses on managing information security risks across an organization. Provides a holistic approach that integrates risk management into the organization's governance structure. | 1. Risk framing<br>2. Risk assessment<br>3. Risk response strategy<br>4. Risk monitoring |
| NIST SP 800-30 (Rev 1) | Is dedicated to IT security risk assessment. It outlines a process to identify, prioritize, and manage IT security risks, specifically aimed at protecting organizational operations and assets. | 1.Risk Assessment Preparation<br>2. Risk Assessment<br>3. Risk Communication and Consultation |
| ISO/IEC 27005 | Provides guidelines for information security risk management. Part of the ISO/IEC 27000 series focusing on aspects of information security management systems (ISMS). | 1. Establishing context<br>2. Risk assessment<br>3. Risk treatment<br>4. Risk monitoring<br>5. Risk review |
| COSO Framework | Geared towards enterprise risk management, integrating IS risks with broader organizational risks and aligning them with strategy and performance. | 1. Internal environment<br>2. Objective setting<br>3. Event identification<br>4. Risk assessment<br>5. Risk response<br>6. Control activities<br>7. Information and communication<br>8. Monitoring |
| COBIT | Designed for IT governance and management, it aligns IT activities with business objectives through a comprehensive set of processes and best practices. | 1. Management objectives<br>2. Performance management<br>3. Risk optimization<br>4. Value delivery |

**Table 2.** Alternative IS risk management frameworks.

| Framework/Method | Description | Key Features |
| --- | --- | --- |
| SQUARE | A methodology created at Carnegie Mellon University for identifying and prioritizing the security requirements of IT systems. | Integrates security requirements early in the system development process, using a nine-step process that includes risk assessment and requirements inspection. |
| Fuzzy Expert System for IS Risk | Developed for use in learning management systems to improve accuracy and effectiveness in information security risk assessment. | Utilizes fuzzy logic to handle uncertainty and imprecision, enhancing traditional risk assessment methods. |
| Security Assessment for CPDPS | Framework developed for assessing security in cyber–physical distribution power systems under various intrusion attacks. | Considers the impact of different attack types on system security, reliability, and stability; introduces security metrics for better risk management. |
| SCADA Systems Cybersecurity | Presents a novel method for assessing cybersecurity risks in SCADA systems using a multi-layered approach. | Involves identifying threats, assessing system vulnerabilities, evaluating threat likelihood, determining potential consequences, and prioritizing risks. |
| Entropy Weight Quantitative Risk Assessment | Applied to industrial control systems to identify critical factors and interdependencies affecting system safety and security. | Combines qualitative and quantitative factors in risk assessment, adaptable to the dynamic nature of industrial control systems. |
| Association Analysis for Industrial Control Systems | Utilizes machine learning and association analysis to identify and evaluate risks in industrial control systems. | Provides a comprehensive method that considers both quantitative and qualitative factors in cybersecurity risk assessment. |
| MAGDM Using Probabilistic Linguistic Term Sets | Proposes a novel approach to multi-attribute group decision making in information security risk management. | Uses probabilistic linguistic term sets to handle imprecise information, improving the accuracy and reliability of group decision-making processes. |
| Decision Support System for Information Security | Combines risk assessment, cost-benefit analysis, and optimization techniques to assist in selecting optimal preventative actions. | Integrates multiple components into a single system, allowing data-driven decisions that align with organizational needs and constraints. |
| Dynamic Security Assessment in Power Systems | Develops an integrated transfer learning method for assessing power system security in the presence of unlearned faults and missing data. | Shows superior performance in handling common issues in power systems, improving traditional methods and existing transfer learning approaches. |

In this scientific paper, the authors present a novel method for assessing cybersecurity risks in Supervisory Control and Data Acquisition (SCADA) systems [42], which is based on a multi-layered approach. The method involves the following steps: a. Identifying and categorizing potential threats to the SCADA system. b. Assessing the vulnerabilities of the system to these threats. c. Evaluating the likelihood of each threat occurring. d. Determining the potential consequences of each threat. e. Calculating the overall risk score for each threat based on its likelihood and potential consequences. f. Prioritizing threats based on their risk scores and implementing appropriate security measures to mitigate the highest-priority threats.

The research paper [43] presents the results of applying the entropy weight quantitative risk assessment method to a typical industrial control system. The method effectively identified the most critical factors contributing to the safety and security of the system, as well as the interdependencies between these factors. This enabled the researchers to prioritize the factors and allocate resources more effectively to address the most pressing risks. The novelty of the method lies in its ability to consider both qualitative and quantitative factors in risk assessment, as well as its adaptability to the dynamic nature of industrial control systems. This combination of factors makes the method more comprehensive and effective than traditional risk assessment methods, which often rely on either qualitative or quantitative factors alone.

The novelty of paper [44] lies in its application of association analysis techniques to the field of cybersecurity risk assessment for industrial control systems. By leveraging the power of machine learning and association analysis, the authors provide a more comprehensive and accurate method for identifying and evaluating risks in industrial control systems. Furthermore, the paper highlights the importance of considering both

quantitative and qualitative factors when assessing cybersecurity risks. This is an essential aspect that has often been overlooked in the existing literature.

In scientific paper [45], the authors propose a novel approach to multi-attribute group decision making (MAGDM) using probabilistic linguistic term sets (PLTS). The authors introduce the concept of PLTS, which is a new method for representing and handling imprecise information in decision-making problems. PLTSs are based on the probabilistic interpretation of linguistic terms and provide a more flexible and robust way to deal with uncertainty in comparison to traditional linguistic term sets. The authors present a step-by-step decision-making process that integrates PLTS in MAGDM. This process includes determining the criteria and alternatives, aggregating the preferences of individual decision makers, and calculating the overall preference of each alternative using the PLTS. The result is a more accurate and reliable group decision that takes into account the preferences and opinions of all decision makers.

The novelty of paper [46] lies in its comprehensive and systematic presentation of the historical development, classification, relationships, and applications of fuzzy sets. By providing a clear understanding of these concepts, the paper serves as a valuable resource for researchers and practitioners working in the field of fuzzy set theory and its applications.

The research work of [47] primarily focuses on the development of a novel risk assessment method that utilizes consistency and probabilistic linguistic preference relations. The study aims to address the limitations of traditional risk assessment methods and improve the accuracy and reliability of risk assessment in various domains. The novel method improves the decision-making process by considering both consistency and preference relations. This leads to better-informed decisions and reduces the chances of errors or biases in the risk assessment process.

The novelty of scientific paper [48] lies in the development of a decision support system that combines risk assessment, cost–benefit analysis, and optimization techniques to assist enterprises in selecting optimal preventative actions for information security. While previous research has focused on individual aspects of information security decision making, this paper presents a comprehensive framework that integrates multiple components into a single unified system. By considering both the costs and benefits of preventative actions, organizations can make data-driven decisions that align with their specific needs and resource constraints. The integration of optimization techniques further enhances the system's ability to identify the most effective combination of actions for risk reduction.

Scientific paper [49] presents a novel approach to assess the dynamic security of power systems in the presence of unlearned faults and missing data. The main results of the study include the development of an integrated transfer learning method that combines two existing transfer learning models for power system fault analysis. The authors demonstrate that their proposed method can effectively handle the challenges posed by unlearned faults and missing data, which are common issues in power systems. They evaluate the performance of their method using various benchmark datasets, showing that it outperforms traditional methods and existing transfer learning approaches in terms of accuracy and efficiency.

## 2. Materials and Methods

The initial step in developing the flexible information security risk model involves an exhaustive study of existing information security standards related to risk assessment. This comprehensive process enables the compilation of an expert-generated list of information security risks. This initial list forms the basis for subsequent expansion and refinement. Conducting a study of information security risk management processes based on the analysis of ISO/IEC 27000, ISO/IEC 27001, ISO/IEC 27002, ISO/IEC 27003, ISO/IEC 27005, ISO/IEC 27014, ISO/IEC 15408, IEC 62443, ETSI EN 303 645, NIST 800-30, ISO/IEC TR 13335-3:1998, BS 7799 standards is necessary for establishing best practices [50]. These standards are recognized globally as best practices for information security (IS) and risk

management. They provide guidelines, frameworks, and methodologies that organizations can follow to enhance their information security posture. Studying these standards helps organizations adopt industry-accepted practices. Many regulatory bodies and industry-specific organizations require adherence to these standards as a condition of doing business. Compliance with these standards may be mandatory for certain industries or organizations that handle sensitive information, such as healthcare, finance, or critical infrastructure. Conducting a study helps organizations ensure they meet these compliance requirements. Information security is critical for protecting an organization's assets, reputation, and customer trust. A study of these standards helps organizations identify and mitigate risks effectively, reducing the likelihood of security incidents, data breaches, and other cyber threats. These standards cover various aspects of information security, including risk management, security controls, incident response, and security governance. Studying them provides a comprehensive understanding of information security, allowing organizations to address security holistically rather than in isolation. Effective communication of security risks and controls within an organization is essential. These standards provide a common language and framework for discussing security risks, making it easier for stakeholders to understand and collaborate on security-related matters.

To enhance the flexibility of the information security risk model, cluster analysis is employed using the k-means algorithm applied to the texts of information security standards. This data-driven approach aims to uncover hidden patterns and relationships within the standards that may not be immediately apparent through manual examination. The utilization of the k-means algorithm can allow for the grouping of similar textual elements, facilitating the identification of clusters representing potential information security risks [51]. The primary objective of this analysis is to broaden the existing list of information security risks by systematically categorizing and organizing related concepts. By doing so, the model can demonstrate its adaptability to evolving knowledge and insights in the field. This approach enables the inclusion of nuanced and emerging risks that might not have been explicitly addressed in traditional, expert-compiled lists. The k-means clustering method was chosen for its ability to handle large datasets and efficiently organize textual information into distinct groups. This facilitates a more nuanced understanding of the relationships between different elements within the information security standards. The expanded list of risks derived from this analysis can contribute to the model's flexibility, ensuring its relevance and applicability in dynamic and evolving information security landscapes. Despite the utilization of cluster analysis on a substantial corpus of standardized text, the methodology predominantly relies on data derived from existing information security standards. These standards may not comprehensively encapsulate all potential risks within a dynamically evolving threat landscape. To mitigate the limitations inherent in depending solely on these standards, it is imperative to integrate a wider array of data sources. This enhancement should include the analysis of recent cybersecurity incident reports, incorporation of real-time threat intelligence feeds, and the application of advanced predictive analytics. Such an expansion is crucial for capturing the rapidly changing nature of cyber threats, which may not be adequately represented in the conventional datasets.

Utilizing machine learning methods, a dendrogram [52] of information security risks will construct to further demonstrate the model's flexibility. Machine learning algorithms are employed to uncover latent patterns and relationships within the data, facilitating the identification of new and emerging risks. This dynamic approach can ensure that the risk model could adapt to changing threat landscapes over time. In order to augment the flexibility of the information security risk model, a heat map of information security risks will be compiled. Heat maps, as a visual representation tool, are employed to illustrate the relative significance and interconnections among the identified risks [53]. This approach aims to provide a dynamic and intuitive means of understanding the complex relationships between different risk factors. The construction of the heat map involves assigning varying degrees of intensity or color gradients to different risk categories based on their relative importance. Darker shades or higher intensities typically represent higher levels of risk

significance, while lighter shades indicate lower levels. This visual representation can allow for a quick and comprehensive assessment of the risk landscape, emphasizing the criticality of certain risk factors over others.

The creation of a thesaurus [54], encompassing definitions, synonyms, and antonyms, plays a crucial role in optimizing the risk assessment process. This optimization can showcase the model's ability to refine and improve risk descriptions, making them more precise and adaptable to changing terminology and language conventions in the field. A pivotal aspect of the model's flexibility will involve the development of an ontological model of information security risks. This ontological structure [55] can provide a standardized and structured framework for classifying risks. By categorizing risks based on their ontological relationships, the model can become more adaptable and scalable, accommodating new risk categories as they emerge.

## 3. Results

### *3.1. Study of Information Security Risk Management Processes Based on Standards Analysis*

We analyzed the following information security standards: ISO/IEC 27000, ISO/IEC 27001, ISO/IEC 27002, ISO/IEC 27003, ISO/IEC 27005, ISO/IEC 27014, ISO/IEC 15408, IEC 62443, ETSI EN 303 645, NIST 800-30, ISO/IEC TR 13335-3:1998, BS 7799. After analyzing all these standards related to information security, we grouped them into 65 types of risk.

ISO/IEC 27000 is not a standard that describes specific information security risks. Instead, it is part of the ISO/IEC 27000 series, which provides guidelines and best practices for establishing, implementing, maintaining, and continually improving an information security management system (ISMS) within the context of an organization.

ISO/IEC 27001 is the main standard that specifies the requirements for establishing, implementing, maintaining, and continually improving an ISMS. It provides a structured approach for identifying, assessing, and treating information security risks. There is a clear requirement in this standard that needs to be mentioned: Information Security Policies and Procedures.

ISO/IEC 27002 provides a comprehensive set of guidelines and best practices for information security controls. While it does not list specific information security risks, it outlines a wide range of controls that organizations can implement to mitigate various types of risks. These controls cover different areas of information security. There is a clear requirement in this standard that needs to be mentioned: Information Security Policies and Procedures; Defining Roles and Responsibilities; Risk of Insider Threats of Data; Risk of Unauthorized Access to Data; Risk of Data Breaches; Risk of Encryption and Cryptographic Controls; Risk of Damage Servers and Data Centers; Risk of Unauthorized Physical Access to Servers and Data Centers; Risk of Theft Servers and Data Centers; Risk of Data Disruptions; Risk of Data Loss; Risk of Unauthorized Changes in Data; Risk of Eavesdropping Through Communication Channels; Risk of Data Interception Through Communication Channels; Risk of Unauthorized Access to Sensitive Information Through Communication Channels; Establishing Security Requirements for Third-Party Suppliers; Risk of Prolonged Disruptions to Operations; Legal Risk; Reputational Risk; Risk of Human Errors Leading to Security Breaches; Monitoring Security Events and Measuring Security Metrics.

ISO/IEC 27003 provides guidelines for the implementation of ISMS based on the requirements specified in ISO/IEC 27001. It focuses on the processes and activities involved in planning, establishing, implementing, operating, monitoring, reviewing, maintaining, and improving the ISMS. While ISO/IEC 27003 does not explicitly list information security risks, it provides guidance on how to structure and manage an ISMS to address various risks effectively. There is a clear requirement in this standard that needs to be mentioned: Risk of Information Confidentiality; Risk of Information Integrity; Risk of Information Availability; Information Security Policies and Procedures; Organization's Specific Risk; Defining Roles and Responsibilities.

ISO/IEC 27005 is a standard that provides guidelines for information security risk management. It outlines a structured and systematic approach to identifying, assessing, and treating information security risks within the context of an organization. While it does not provide an exhaustive list of specific information security risks, it offers a framework for organizations to assess and manage risks based on their unique circumstances. There is a clear requirement in this standard that needs to be mentioned: Risk of Information Confidentiality; Risk of Information Integrity; Risk of Information Availability; Risk to Potential Business Impact; Legal Risk; Regulatory Requirements Risk; Stakeholder Concerns; Residual Risk, which is the Risk That Remains After Implementing Controls.

ISO/IEC 27014 is a standard that provides guidelines for information security governance. It focuses on establishing and maintaining a framework for governing an organization's information security activities and aligning them with overall business objectives. While it does not explicitly list specific information security risks, it provides guidance on how to manage information security within a governance context. There is a clear requirement in this standard that needs to be mentioned: Risk of Inadequate Resource Allocation; Risk of Lack of Buy-In by Senior Management; Risk of Human Errors Leading to Security Breaches; Risk of Inadequate Information Security Measures Associated with Proper Allocation of Budget and Personnel; Risk of Non-Compliance and Inadequate Security Practices; Legal Risk; Regulatory Requirements Risk.

ISO/IEC 15408, commonly known as the Common Criteria (CC), is a standard that specifies a framework for evaluating the security properties of information technology products and systems. It provides a structured approach to evaluating the security features and capabilities of products and systems. While the CC does not explicitly list specific information security risks, it outlines the methodology and criteria for assessing and certifying the security of these products and systems. The methodology and criteria outlined in the standard contribute to effective risk management in the context of evaluating and certifying the security of information technology products and systems.

IEC 62443 is a series of standards developed by the International Electrotechnical Commission that addresses the security of industrial automation and control systems, often referred to as industrial cybersecurity. These standards provide guidelines for establishing a systematic and comprehensive approach to managing the cybersecurity risks associated with industrial control systems (ICS). While IEC 62443 does not explicitly list specific information security risks, it defines processes and measures to mitigate risks in industrial environments. There is a clear requirement in this standard that needs to be mentioned: Risk of Including Access Control; Risk of Authentication; Risk of Encryption and Cryptographic Controls; Risk Associated with Third-Party Vendors; Risk Associated with Partners Involved in the Industrial Ecosystem; Risk Associated with External Connections.

ETSI EN 303 645 specifically focuses on cybersecurity for IoT devices. There are common information security risks that are discussed in such a standard: Risk of Automated Attacks Against a Class or Type of Device; Licensing Risk; Risk of Authentication; Risk of Authorization; Risk of Insecure or Default Credentials; Risk of Unauthorized Access to IoT Devices; Lack of Secure Update Mechanisms; Risk of Connections to Insecure Network; Risk of Data Interception Through Communication Channels; Risk of Unauthorized Access to Sensitive Information Through Communication Channels; Risk of Encryption and Cryptographic Controls; Risk of Data Breaches; Risk of Vulnerable Interfaces; Risk of Vulnerable Interfaces APIs; Risk of Insufficient Physical Security; Lack of Privacy Controls; Risk of Insecure Software/Firmware; Risk of Weak Device Management; Denial of Service (DoS) Risk; Risk of Insecure Default Settings; Risk of Insecure Mobile Application Interfaces; Lack of Security Monitoring; Lack of Logging; Risk Appearing When Writing Code; Risk Associated with Third-Party Vendors; Device Integration Risk.

NIST SP 800-30 is a guidance document from the National Institute of Standards and Technology (NIST) that provides guidance on risk assessment processes for information technology systems. It outlines various steps and considerations for conducting risk assessments. It does not provide a list of specific risks itself, but rather a framework for

identifying and assessing risks. There are common information security risks that are discussed in such a standard: Risk of Physical Loss or Theft of Critical Assets, Hardware, Software, Data; Risk of Virtual Loss or Theft of Critical Assets, Hardware, Software, Data; Risk of Unauthorized Access; Risk of Data Breaches; Malware Risk; Viruses Risk; Risk of Insider Threats of Data; DoS Risk; Phishing Risk; Social Engineering Risk; Risk of Authentication; Risk of Authorization; Risk of Connections to Insecure Network; Lack of Secure Update Mechanisms; Risk of Insufficient Physical Security; Risk Associated with Third-Party Vendors; Legal Risk; Regulatory Requirements Risk; Risk of Information Availability; Environmental Risk.

ISO/IEC TR 13335-3 provides guidelines for the management of information security risk. There are common information security risks that are discussed in such a standard: Risk of Unauthorized Access to Data; Risk of Unauthorized Physical Access to Servers and Data Centers; Risk of Authentication; Risk of Authorization; Risk of Data Breaches; Malware Risk; Viruses Risk; Risk of Insider Threats of Data; Social Engineering Risk; Risk of Connections to Insecure Network; Risk of Encryption and Cryptographic Controls; DoS Risk; Risk Associated with Third-Party Vendors; Lack of Secure Update Mechanisms; Information Security Policies and Procedures; Legal Risk; Regulatory Requirements Risk; Risk of Information Availability; Risk of Data Breaches; Risk of Information Confidentiality.

BS 7799 is a former British standard that provided guidelines for ISMS. It was later adopted as the international standard ISO/IEC 27001. As such, the information security risks discussed in BS 7799 are similar to those in ISO/IEC 27001. There are common information security risks that are discussed in such a standard: Risk of Unauthorized Access to Data; Risk of Unauthorized Physical Access to Servers and Data Centers; Risk of Authentication; Risk of Authorization; Risk of Data Breaches; Malware Risk; Viruses Risk; Risk of Insider Threats of Data; Social Engineering Risk; Risk of Connections to Insecure Network; Risk of Encryption and Cryptographic Controls; DoS Risk; Risk Associated with Third-Party Vendors; Lack of Secure Update Mechanisms; Information Security Policies and Procedures; Legal Risk; Regulatory Requirements Risk; Risk of Information Availability; Risk of Data Breaches.

### *3.2. Data Collection and Preprocessing*

We collected a dataset of text documents from IS standards that we want to analyze. We meticulously gathered all the textual content from the studied standards that references the term “risk”. The collected text consists of 2325 words, 13,672 characters without spaces, and the word “risk” occurs 746 times. Then we cleaned the collected text data, removing unnecessary characters, punctuation marks, and special characters. Cleaning text data is an important step in preparing data for analysis or natural language processing (NLP) tasks. We converted the entire text to lowercase to ensure consistency in text data. This prevents variations in capitalization from being treated as distinct words. Punctuation marks like commas, periods, exclamation marks, and question marks are often not relevant for many NLP tasks. Extra whitespace characters (such as tabs and multiple spaces) need to be removed or replaced with single spaces. Stop words are common words like “and”, “the”, “is”, etc., that might not add much meaning to analysis. Depending on our analysis, we removed them using libraries like Natural Language Toolkit (NLTK) or spaCy. The purpose of this function is to clean and preprocess the input text by applying a series of text processing operations. These operations include the following:

- Converting the text to lowercase;
- Removing punctuation characters;
- Removing numeric digits (numbers);
- Removing special characters (non-alphanumeric and non-whitespace characters);
- Removing extra whitespace;
- Removing common English stopwords.

Then, we performed word tokenization and subword tokenization using techniques like Byte-Pair Encoding (BPE) or WordPiece. Tokenization is the process of breaking

down a text into individual units, which are usually words or subword units, for further analysis. Word tokenization involves splitting the text into individual words. The NLTK and spaCy libraries are commonly used for this purpose. Byte-Pair Encoding is a subword tokenization technique that breaks down words into smaller subword units. WordPiece is another subword tokenization technique, primarily used by the BERT model.

### 3.3. Data Collection and Preprocessing

For Feature Extraction, we chose the method Term Frequency–Inverse Document Frequency (TF-IDF). Converting text into numerical features using TF-IDF is a common technique in NLP for machine learning. We imported the necessary Python libraries, including scikit-learn for TF-IDF vectorization. Then, we created a TF-IDF vectorizer object and used it to convert our text data into numerical features.

Each row in the resulting matrix represents a document, and each column represents a unique word (term) from the entire corpus. The values in the matrix represent the importance of each word within each document relative to its occurrence across all documents. Higher values indicate greater importance. We used the TF-IDF matrix as input features for machine learning algorithms. Typically, this matrix is used for tasks like text classification, clustering, or any other NLP-related task. The TF-IDF vectorization process captures the importance of words within individual documents relative to the entire corpus. It is a common technique for text-based machine learning tasks, helping algorithms understand the significance of words in different documents while considering their frequency across the entire dataset. Applying dimensionality reduction techniques like Principal Component Analysis (PCA) or t-Distributed Stochastic Neighbor Embedding (t-SNE) involves several steps to reduce the dimensions of data while preserving important information. To apply Principal Component Analysis (PCA) to a dataset represented as a TF-IDF array, we need to use Python and relevant libraries like NumPy and scikit-learn. First, we need to load the TF-IDF data from JSON file into a suitable data structure like a NumPy array or a DataFrame. To reduce the dimensionality of TF-IDF data, we applied PCA. We used the scikit-learn library for this purpose sklearn.decomposition import PCA. After applying PCA, we analyzed the explained variance to understand how much information is retained in the reduced dimensions. This can help us decide on the appropriate number of components.

### 3.4. Calculation of the Values of the Euclidean Distance between the Risks of IS

Analyzing the cluster assignments and exploring the characteristics of each cluster is an important step in understanding the results of a k-means clustering algorithm. Cluster centroids represent the "average" or "center" of each cluster in the feature space. These centroids are vectors with the same dimensions as original data. We accessed the cluster centroids in scikit-learn after fitting the k-means model. The centroids can provide insights into the central tendencies of each cluster. To identify the most representative documents within each cluster, we calculated the distance between each document and the cluster centroid. Documents that are closest to the centroid are typically considered the most representative of that cluster. We can then access the documents from the original data using these indices to examine what makes each cluster unique. We used the Euclidean distance on data. The Euclidean distance $d$ of two data cases $(x_1, x_2)$ is defined as the square root of the sum of squared differences $d(x, y) = \sqrt{\sum |x_i - y_i|^2}$. The Euclidean metric has a great advantage in that it remains unchanged even when a common value is added to each variable of the data. This means that any translations or shifts in the data will not affect the metric. However, it is worth noting that the Euclidean distance is not scale invariant. This means that if the data were to be multiplied by a common factor, it would change the distance calculations. When it comes to class or cluster problems, using the squaring function in Euclidean distance has a significant implication. It effectively sets the limit for inner class distances at 1 and between (inter) class distances. This implicit definition helps in determining the proximity and relationships within and between different classes, making

it a valuable tool in problem solving. Although it makes sense to use this approach for straightforward distributions like normal (Gaussian) distributions, it may not be as suitable for more intricate variables with bimodal or multimodal distributions. Other methods may be more appropriate in such cases to ensure accurate analysis and interpretation of the data. When embarking on the exciting journey of data mining and knowledge discovery in multivariate data, it is crucial to start by thoroughly analyzing the distribution of individual variables. This foundational step provides valuable insights that form the building blocks for uncovering meaningful patterns and relationships within data.

### *3.5. Application of the K-Means Cluster Analysis Method to Create Groups of Similar IS Risks (Clusters)*

Clustering with k-means is a popular unsupervised machine learning technique used to group similar data points into clusters. Our data are in a format where each row represents a data point, and each column represents a feature. K-means is a numerical clustering algorithm and may not work well with categorical data without proper preprocessing. One of the critical steps in k-means is determining the number of clusters (K). We can use methods like the Elbow Method or Silhouette Score to help us choose an appropriate value for K. Figure 1 displays the Silhouette Score.

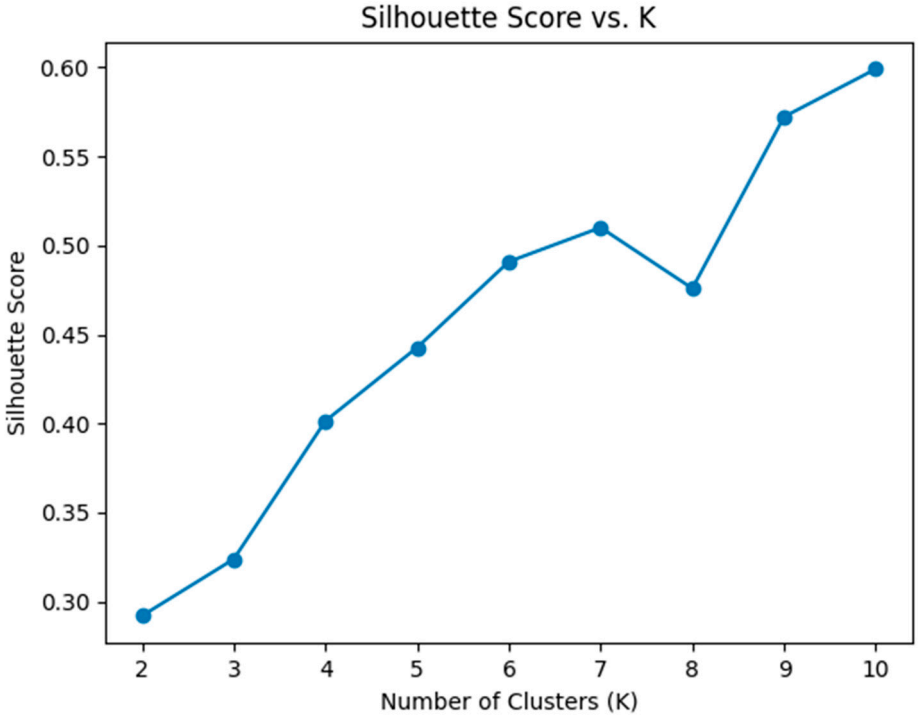


**Figure 1.** Silhouette Score of clusters.

Based on the plot, we chose the value of K where the Silhouette Score starts to plateau or peaks. This is often referred to as the "elbow" of the curve. Our Silhouette values are above average. Actually, the plot says that we have eight clusters. The range between 0.45 and 0.55 is very interesting: our line does not approach a smooth line. A non-smooth line of any clustering criterion, even if it is not horizontal, should be interpreted as "the presence of clear clusters". Once we have chosen the number of clusters (K), we can fit the k-means model to data. The visualization of clusters is described in Figure 2.

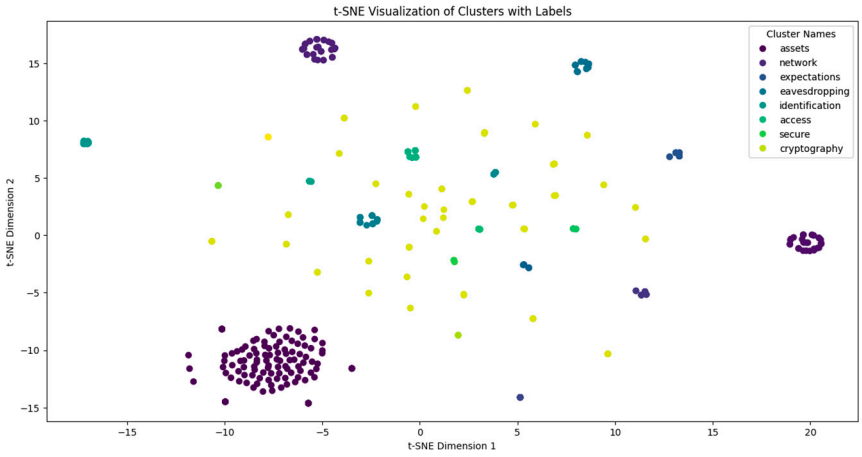


**Figure 2.** Visualization of IS risk clusters.

This t-SNE (t-Distributed Stochastic Neighbor Embedding) visualization presents a two-dimensional representation of clusters related to information security risks. Each cluster is labeled and colored distinctly according to different types of risks such as assets, network, expectations, eavesdropping, identification, access, secure, and cryptography. Well-separated clusters that correspond to known categories or expected results based on domain knowledge indicate good model performance. Overlap or mixed clusters might suggest either mixed data characteristics or a need for tuning t-SNE parameters. For example, clusters labeled as "assets" and "cryptography" suggest that the dataset contains distinct groups where each group shares common characteristics about these specific types of risks. This visualization helps in identifying relationships between different risk types, understanding the distribution and density of the data, and identifying outliers or anomalies. Clusters can indicate areas where security resources need to be focused or where further analysis is required. This in-depth statistical analysis and explanation of the t-SNE visualization provides a solid foundation for understanding the underlying data structure, thereby enhancing the credibility and utility of the clustering in practical scenarios. This approach helps ensure that the insights derived from the visualization are based on sound statistical principles.

Based on the results of information security risks clustering, we received eight distinct clusters: assets, network, expectations, eavesdropping, identification, access, secure, and cryptology. Performing hierarchical clustering with tokenized data and a TF-IDF matrix in JSON format involves several steps. Hierarchical clustering is a method for grouping similar data points into clusters based on their similarity or distance. We used a hierarchical clustering algorithm like agglomerative clustering to create a hierarchical clustering tree (dendrogram). To visualize the clusters, we chose a suitable threshold from the dendrogram to cut the tree and form clusters. This threshold depends on the problem and the level of granularity we want. The results of constructing a hierarchical dendrogram are described in Figure 3.

Ensuring that the chosen distance metric appropriately reflects the similarities or differences in the data is critical for meaningful clusters. Standardizing or normalizing data can be crucial when variables have different scales, as it prevents variables with larger scales from dominating the distance calculations. Each leaf on the dendrogram represents an individual data point (in your case, a type of risk), and the height at which branches merge represents the distance at which clusters are merged. Close merging indicates high similarity. Based on the results of constructing the IS risks dendrogram, we identified 16 explicit clusters. Two clusters "assets" and "access" are present both in the dendrogram and in the cluster visualization. The following clusters are also present in the dendrogram: risk, security, information, controls, assets, organizational, threat, assessment, device, policy, system, access, likelihood, management, loss, operational, unauthorized, cryptography, error, network, authentication, integrity, confidentiality, identification, theft, requirements,

communication, document, availability, stakeholders, regulatory, legal, misuse, fraud, technology, environmental, operations, leakage, data, secure, assessments, climate, environment, operation, testing, log, DOS, encryption, verification, licensing, connection, failures, equipment, external, communicate, channels, exploitation, injection, validation, wireless, commerce, eavesdropping, requirement, and expectations.

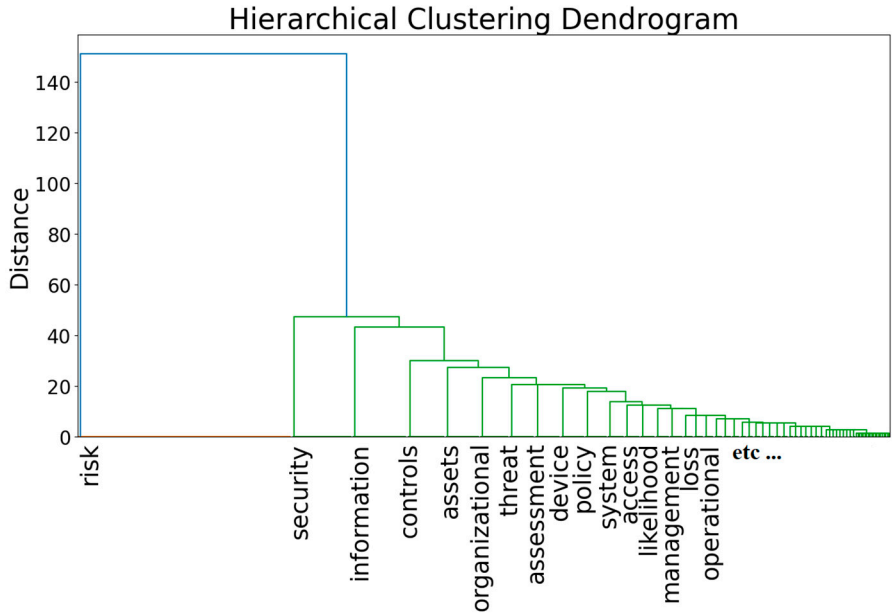


**Figure 3.** Dendrogram of IS risk clusters.

To ensure the completeness of all IS risks, we built a heat map of the clusters. This will allow us to more accurately understand the importance of each cluster for further building a flexible information security risks assessment model. Creating an NLP heat map with tokenized data and a TF-IDF matrix involves visualizing the TF-IDF values for different tokens and documents. The results of building a heat map are described in Figure 4.

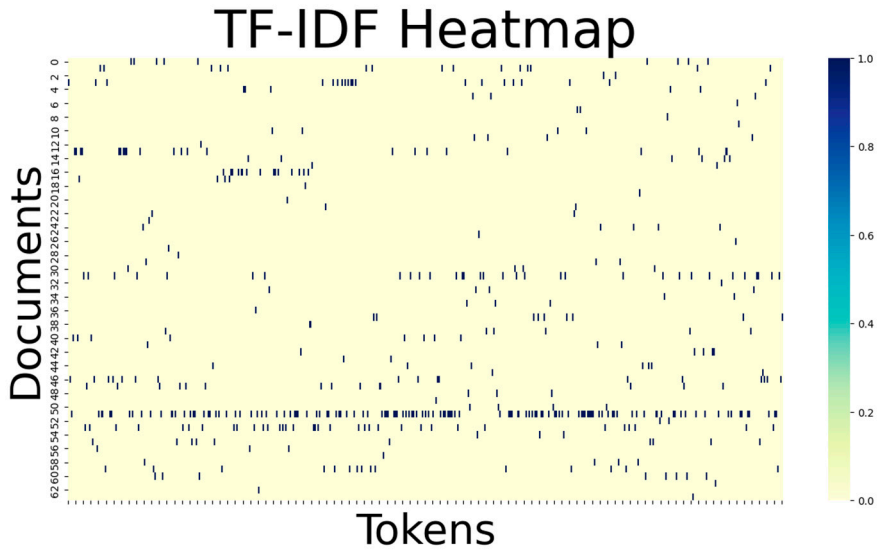


**Figure 4.** Heat map of IS risk clusters.

This heat map provides a comprehensive and visually interpretable way to analyze text data, revealing patterns and key terms across documents that might not be immediately apparent from raw text alone. By ensuring the statistical rigor of TF-IDF calculations and thoughtful visualization, the heat map is a powerful tool for text analysis in information

security risk management. Each row represents a document, and each column represents a token. This layout helps to identify which tokens are key in which documents. The product of TF and IDF scores of a term is high when a term is frequent in a small number of documents, thus lending high discriminative power to the term within the corpus. The color scale (from 0 to 1) represents the TF-IDF score, where darker colors indicate higher importance of the token in the corresponding document. The results of the IS risk cluster heat map show a similar result to the cluster dendrogram, some new clusters were found. The following clusters are present in the heat map of IS risk clusters: vendors, interfaces, integration, viruses, malware, phishing, social, risk, security, information, controls, assets, organizational, threat, assessment, device, policy, system, access, likelihood, management, loss, operational, unauthorized, cryptography, error, network, authentication, integrity, confidentiality, identification, theft, requirements, communication, document, availability, stakeholders, regulatory, legal, misuse, fraud, technology, environmental, operations, leakage, data, secure, assessments, climate, environment, operation, testing, log, DOS, encryption, verification, licensing, connection, failures, equipment, external, communicate, channels, exploitation, injection, validation, wireless, commerce, eavesdropping, requirement, and expectations.

### *3.6. Drawing up a Thesaurus for Assessing Information Security Risks*

Compiling a thesaurus of information security risks is essential for helping organizations identify and categorize potential information security risks that they may face. By having a structured and comprehensive list, organizations can ensure that they do not overlook any significant threats. A thesaurus of information security risks aids in the assessment of the potential impact and likelihood of each risk. This information is crucial for prioritizing risks and allocating resources effectively to mitigate them. Also, it provides a common language and terminology for discussing security risks within an organization. This facilitates communication among various stakeholders, such as IT teams, management, and employees, ensuring a shared understanding of the risks. A comprehensive list of risks can be integrated into an organization's risk management framework, helping in the development of risk assessment methodologies and risk treatment plans. It often happens that in many standards, the same type of information security risk is called by different words. A thesaurus is necessary to determine synonyms and antonyms and determine hidden relationships (dependencies) between entities. Therefore, compiling a thesaurus will allow us to optimize the structural organization of all entities in the subject area before we begin to develop an information security risk assessment model. When compiling the thesaurus, we used the results obtained by expert analysis and machine learning methods. In total, we received 91 information security risks.

After analyzing the compiled thesaurus, we decided not to use the following types of risks in our study: Risk of Including Access Control; Risk Associated with Communicate; Validation Risk; Risk of Failures. The terms "Risk of Including Access Control", "Risk Associated with Communicate", and "Validation Risk" in the context of information security risks are not usually used as a separate concept. The term "Risk of Failures" encompasses a wide range of concepts as it pertains to the potential vulnerabilities and dangers that are connected to the possibility of failures or breakdowns in an organization's information technology systems, processes, or controls. Failures can arise due to a range of factors, such as technical complications, human mistakes, or external threats. These failures have the potential to cause disruptions in operations, loss of data, or breaches in security.

Checking for synonyms allowed us to exclude the following risks from the thesaurus: Risk of Data Loss (accepted analogue is "Risk of Information Integrity"); Organizational Risk (accepted analogue is "Operational Risk"); Risk Associated with Equipment (accepted analogue is "Risk of Physical Loss or Theft of Critical Assets, Hardware, Software, Data"); External Risk (accepted analogues are "Environmental Risk" and "Climate Risk"); Exploitation Risk (accepted analogues are "Risk of Vulnerable Interfaces" and "Risk of Vulnerable Interfaces APIs"); Injection Risk (accepted analogue is "Risk Appearing When

Writing Code"); Risk Related with Wireless (accepted analogues are "Risk of Eavesdropping Through Communication Channels", "Risk of Data Interception Through Communication Channels", and "Risk of Unauthorized Access to Sensitive Information Through Communication Channels"); Commerce Risk (accepted analogue is "Risk to Potential Business Impact"). After a detailed analysis of the thesaurus, we are left with 79 types of information security risks.

### *3.7. Construction of an Ontological Model for Assessing Information Security Risks*

Building IS risks ontology is highly relevant and beneficial in the field of cybersecurity and information technology for providing a structured and standardized way to represent knowledge about information security risks. It defines the key concepts, relationships, and attributes relevant to this domain, making it easier for security professionals, researchers, and decision makers to understand and work with this information. An ontology for information security risks can help organizations assess and manage these risks more comprehensively. It can facilitate the identification of vulnerabilities, threats, and potential consequences, leading to more informed risk mitigation strategies. Also, ontology can help organizations map their security practices and risk management processes to regulatory requirements, ensuring compliance and reducing legal risks. Creating "IS risks" ontology based on the created thesaurus is a practical and logical approach for providing a standardized and accepted set of terms and synonyms related to a specific domain. When creating an ontology, it is essential to use standardized terminology to ensure clarity and consistency in how information security risks are described and categorized. Relying on a thesaurus helps maintain this consistency. By basing an ontology on an established thesaurus, it ensures that the concepts and terms used in the ontology align with the common understanding and language used by professionals in information security. This makes the ontology more accessible and relevant to practitioners. When constructing the ontology, we combined all 67 risks into 14 groups for better understanding and classification of information security risks. These information security risk groups are displayed in Figure 5.

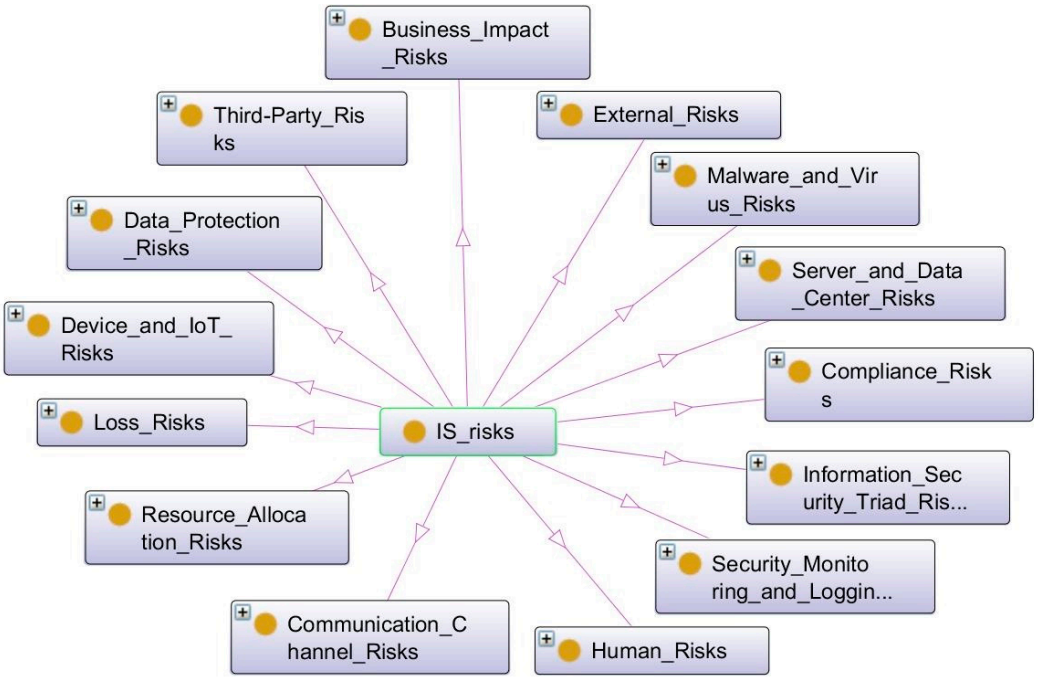


**Figure 5.** Information security risk groups.

Risks are divided into the following groups: Compliance Risks; Data Protection Risks; Server and Data Center Risks; Loss Risks; Communication Channel Risks; Information Security Triad Risks; Business Impact Risks; Resource Allocation Risks; Third-Party Risks;

Device and IoT Risks; Security Monitoring and Logging Risks; Malware and Virus Risks; Human Risks; External Risks.

The Communication Channel Risks group includes the following risks: Risk of Eavesdropping Through Communication Channels; Risk of Data Interception Through Communication Channels; Risk of Unauthorized Access to Sensitive Information Through Communication Channels.

The Business Impact Risks group includes the following risks: Risk to Potential Business Impact; Residual Risk (Remaining Risk After Implementing Controls); Reputational Risk; Organization's Specific Risk; Operational Risk; Risk of Prolonged Disruptions to Operations.

The Compliance Risks group includes the following risks: Information Security Policies and Procedures; Risk Related with Document; Defining Roles and Responsibilities; Legal Risk; Regulatory Requirements Risk; Risk of Non-Compliance and Inadequate Security Practices; Risk of Expectations; Risk of Authentication; Risk of Authorization; Risk of Unauthorized Access to Data; Risk of Unauthorized Changes in Data; Risk of Identification.

The Data Protection Risks group includes the following risks: Risk of Data Breaches; Risk of Encryption and Cryptographic Controls; Cryptography Risk.

The Device and IoT Risks group includes the following risks: Risk of Automated Attacks Against a Class or Type of Device; Licensing Risk; Risk of Insecure or Default Credentials; Risk of Unauthorized Access to IoT Devices; Lack of Secure Update Mechanisms; Risk of Connections to Insecure Network; Risk of Vulnerable Interfaces; Risk of Vulnerable Interfaces APIs; Risk of Insecure Mobile Application Interfaces; Risk of Insufficient Physical Security; Lack of Privacy Controls; Risk of Insecure Software/Firmware; Risk of Weak Device Management; Device Integration Risk; Risk of Insecure Default Settings; Technology Risk.

The External Risks group includes the following risks: Environmental Risk; Climate Risk.

The Information Security Triad Risks group includes the following risks: Risk of Information Confidentiality; Risk of Information Integrity; Risk of Information Availability.

The Human Risks group includes the following risks: Phishing Risk; Social Engineering Risk; Risk of Misuse; Risk of Fraud; Risk of Human Errors Leading to Security Breaches; Risk of Insider Threats of Data; Risk of Error; Risk Appearing When Writing Code.

The Malware and Virus Risks group includes the following risks: Malware Risk; Viruses Risk.

The Loss Risks group includes the following risks: Risk of Data Disruptions; Risk of Leakage; Risk of Physical Loss or Theft of Critical Assets, Hardware, Software, Data; Risk of Virtual Loss or Theft of Critical Assets, Hardware, Software, Data.

The Resource Allocation Risks group includes the following risks: Risk of Inadequate Resource Allocation; Risk of Lack of Buy-In by Senior Management; Risk of Inadequate IS Measures Associated with Proper Allocation of Budget and Personnel.

The Third-Party Risks group includes the following risks: Risk Associated with Third-Party Vendors; Risk Associated with Partners Involved in the Industrial Ecosystem; Risk Associated with External Connections; Establishing Security Requirements for Third-Party Suppliers; Risk Associated with Stakeholders; Stakeholder Concerns.

The Server and Data Center Risks group includes the following risks: Risk of Damaged Servers and Data Centers; Risk of Unauthorized Physical Access to Servers and Data Centers; Risk of Theft of Servers and Data Centers; Denial of Service (DoS) Risk; System Risk.

The Security Monitoring and Logging Risks group includes the following risks: Lack of Security Monitoring; Lack of Logging; Monitoring Security Events and Measuring Security Metrics; Testing Risk; Verification Risk; Likelihood Risk.

To facilitate the practical application of the thesaurus and ontology model, a structured guide for organizations is delineated. To ensure the thesaurus remains up-to-date and reflective of current cybersecurity landscapes, organizations are advised to establish a sys-

tematic updating schedule. The frequency of updates—whether quarterly, semi-annually, or annually—should align with the organization's specific exposure to evolving risks and technological advancements. A designated individual or team should be tasked with the maintenance of the thesaurus. This responsibility includes the continuous monitoring of emerging risks, the periodic updating of terminologies, and the integration of user feedback. To streamline this process, it is recommended that organizations implement a structured feedback mechanism. This could take the form of regular user surveys or a dedicated digital feedback platform, facilitating ongoing improvements to the thesaurus. Updates to the thesaurus should be prompted by specific triggers such as the adoption of new technologies within the organization, which may introduce novel security risks. Furthermore, any security incident should prompt an immediate review of the thesaurus to determine if it encompasses the newly encountered risks. Similarly, modifications in information security laws and regulations should be closely monitored to ensure the thesaurus accurately reflects compliance-related risks, thus maintaining its relevance and utility in dynamic regulatory environments.

To enhance the management of the thesaurus and ontology within organizations, it is advisable to employ specific tools and software designed to automate and simplify updates. Prominent examples include Protégé and Apache Jena, which necessitate proficiency in ontology management. Therefore, the recruitment or development of personnel skilled in these areas is crucial. Protégé is available at no cost, whereas tools such as PoolParty and OntoText GraphDB may incur subscription fees or licensing costs, necessitating budget considerations. For organizations lacking the internal capacity to handle these tools, outsourcing to firms that specialize in semantic technologies and data management represents a viable and cost-effective solution. These firms not only provide the requisite expertise but also offer ongoing support essential for managing sophisticated ontological systems.

In subsequent studies, we plan to employ the ontology and thesaurus of information security risks to devise a novel fuzzy expert system. This system aims to facilitate rapid and straightforward risk assessments of the information systems utilized within an organization. This approach seeks to enhance the efficiency and accuracy of risk evaluation processes, leveraging the structured knowledge embedded in the ontology and thesaurus to support decision making in cybersecurity management.

## 4. Discussion

The development of a flexible information security risk model is crucial in contemporary times due to the rapidly evolving landscape of information technology and the corresponding escalation of cyber threats. A new flexible information security risk model is shown in Figure 6. Cyber threats are continually evolving and becoming more sophisticated. A static or rigid information security risk model may not adequately adapt to these evolving threats. A flexible model can accommodate new threat vectors, attack methodologies, and vulnerabilities, ensuring a proactive and adaptive approach to risk assessment. Organizations operate in diverse and dynamic environments, each with its own unique set of challenges and risks. A flexible risk model can be tailored to suit specific organizational contexts, considering industry-specific regulations, business processes, technologies, and threat landscapes. This adaptability allows for a more accurate and targeted risk assessment. Legal and regulatory frameworks governing information security are continuously evolving to keep pace with the changing threat landscape. A flexible risk model can incorporate these regulatory requirements and compliance standards, assisting organizations in meeting legal obligations and minimizing legal risks. As data become a critical asset for organizations, a flexible risk model can adopt a data-centric approach, focusing on assessing risks associated with the confidentiality, integrity, and availability of sensitive information. This enables organizations to prioritize and protect their most valuable assets effectively. By providing a tailored and dynamic approach to risk assessment, a flexible risk model helps organizations allocate their resources (financial, human, and technological)

efficiently. This ensures that resources are invested in areas with the highest risk impact, enhancing the overall security posture while optimizing resource utilization.

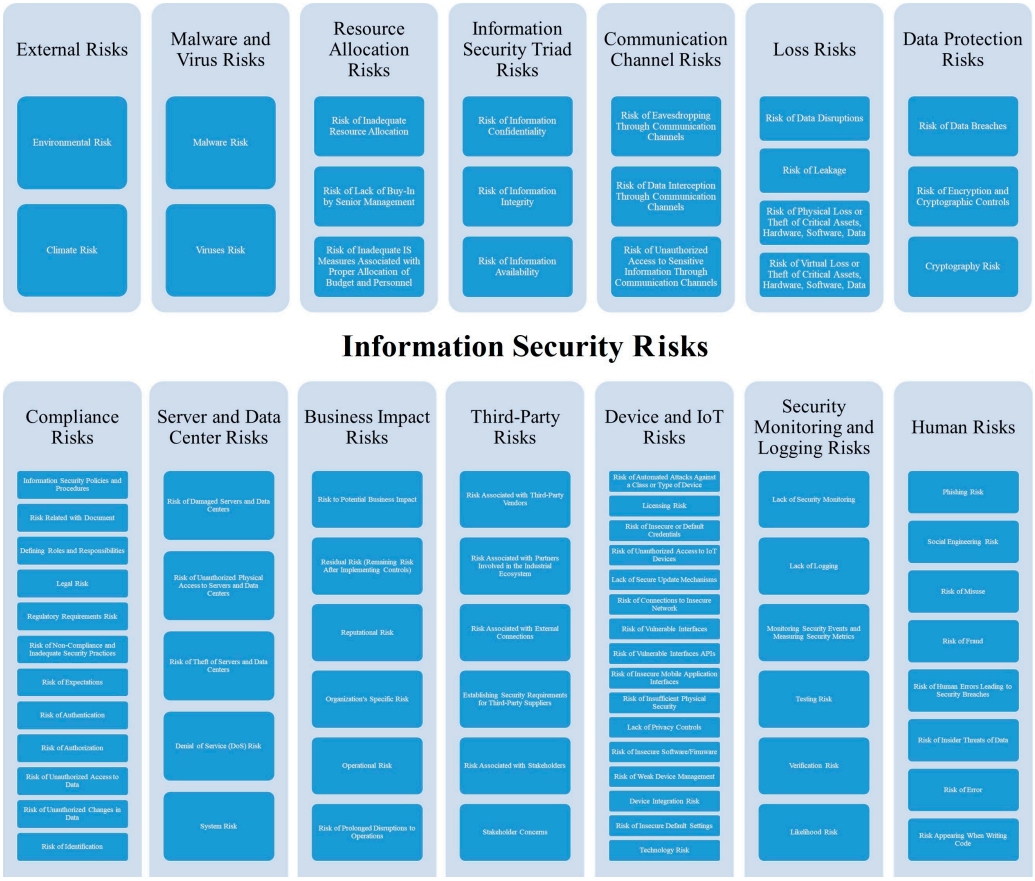


**Figure 6.** New flexible information security risk model.

The proof of flexibility for the new information security risk assessment model can be demonstrated through the iterative and multidimensional approach undertaken during its development, incorporating various methodologies and data sources. Initially, a comprehensive study of all known information security standards was conducted, and an expert list of information security risks was compiled based on this analysis. Texts of all information security standards were subjected to cluster analysis using k-means. This approach helped in identifying patterns and similarities within the standards, thus expanding the list of information security risks based on clustered findings. Machine learning methods were employed to construct an information security risk dendrogram. This approach further expanded the list of risks by uncovering relationships and hierarchies among the risks, enhancing the model's depth and coverage. A heat map of information security risks was compiled, providing a visual representation of risk relationships and intensities. Analyzing this heat map allowed for the identification and inclusion of additional risks, thus enriching the risk list. A thesaurus was compiled, incorporating definitions, synonyms, and antonyms related to information security risks. This thesaurus helped optimize the risk list by standardizing terminology and refining risk descriptions, contributing to a more precise and comprehensive assessment. An ontological model of information security risks was developed, enabling a structured and organized representation of risks. This model facilitated risk classification, providing a basis for further developing a flexible and adaptable risk assessment model. By employing a variety of analytical and data-driven techniques, including expert analysis, cluster analysis, machine learning, heat map analysis, thesaurus compilation, and ontological modeling, the information security risk assessment model was continuously enriched and refined. These iterative processes demonstrate the

model's flexibility, adaptability, and capacity to expand and evolve the list of information security risks. The integration of diverse methodologies and data sources enhances the model's ability to keep pace with evolving threats and organizational contexts, making it a valuable tool for modern information security risk assessment.

Table 3 details the specific features and components of the newly developed information security risk model. It outlines each major component of the model, describes its function, and explains how it contributes to the model's flexibility and effectiveness.

**Table 3.** Features and components of the proposed risk model.

| Component | Description | Role in Risk Assessment | Impact on Model Flexibility |
| --- | --- | --- | --- |
| Machine learning | Utilizes algorithms to identify patterns and make predictions based on data. | Enhances predictive accuracy and identifies emerging threats. | Allows dynamic adaptation to new data and trends. |
| Ontological model | A structured representation defining the types and interrelationships of risk factors. | Facilitates consistent and comprehensive risk categorization. | Enables scalability and integration of new risk categories. |
| Thesaurus | A repository of terms and their synonyms used in risk descriptions to ensure consistency. | Improves the precision of risk descriptions and aids in standardized communication. | Enhances adaptability by incorporating evolving terminologies. |
| Heat maps | Visual tools that display risk intensity across different areas, facilitating easier interpretation of data. | Aids in visualizing and prioritizing risks based on severity and interconnectedness. | Improves decision-making efficiency by highlighting critical areas needing attention. |
| Dendrogram | Tree-like diagrams that show the arrangement of clustered data based on similarity levels. | Helps in understanding the hierarchical relationships among various risk factors. | Supports detailed analysis and refinement of risk clusters for better risk segmentation. |

The development and application of the proposed information security risk model are subject to several limitations that warrant consideration. Primarily, the effectiveness of the model is contingent upon the quality and diversity of the underlying data. In instances where the data are neither comprehensive nor sufficiently varied, the model's capacity to generalize and accurately predict risks across different contexts could be impeded. This limitation underscores the necessity for robust data collection and preprocessing to ensure the representativeness and breadth of the data utilized. Furthermore, the model's architecture integrates multiple complex methodologies, including machine learning, ontological modeling, and cluster analysis. While this integration is designed to enhance the model's analytical capabilities, it also introduces complexity that could hinder its accessibility and usability. Users lacking in technical expertise may find the model particularly challenging to understand and apply effectively, thereby limiting its practicality for a broader audience. Another significant limitation pertains to the testing of the model's performance and scalability. The current evaluation of the model has been confined to controlled experimental or theoretical scenarios, which do not comprehensively simulate real-world operational conditions. Consequently, the model's applicability and effectiveness in large-scale or diverse operational environments remain largely untested. This limitation suggests a potential gap between the model's theoretical performance and its practical utility, indicating the need for further testing and validation under varied real-world conditions to ascertain its scalability and effectiveness across different organizational contexts. These limitations highlight critical areas for future research and development, emphasizing the need for extensive validation studies and the potential refinement of the model to enhance its applicability and usability in diverse settings.

For future research directions, it is imperative to address the limitations of relying exclusively on established standards. Enhancing the data foundation by incorporating a broader spectrum of data sources is critical. Such augmentation should include the systematic analysis of recent cybersecurity incident reports, integration of real-time threat intelligence feeds, and utilization of advanced predictive analytics. This comprehensive

expansion is essential to accurately reflect the dynamic and evolving nature of cyber threats, which traditional datasets may fail to capture adequately.

In subsequent research, leveraging the established ontology of information security risks, we aim to develop a fuzzy expert system tailored for assessing the information security posture of enterprises. This initiative is anticipated to mitigate the integration complexities associated with machine learning, multi-layer clustering, and ontological models. We hypothesize that the introduction of a fuzzy expert system will address the implementation challenges currently faced by small and medium-sized enterprises (SMEs) or organizations lacking extensive technical resources. Moreover, this system is expected to streamline both the implementation and maintenance processes, thereby enhancing usability and accessibility in diverse organizational contexts.

Although ethical considerations are often not extensively addressed in technical cybersecurity research, in future studies, we will contemplate the implications associated with data privacy during the processing of large datasets of organizational risk data. This consideration is critical to ensuring compliance with the evolving data protection regulations and maintaining stakeholder trust in cybersecurity practices.

## 5. Conclusions

In conclusion, this research significantly advances information security risk assessment by integrating established standards with expert knowledge, machine learning, and ontological modeling. This approach has developed a robust foundation that enhances the understanding of the risk landscape. The use of cluster analysis and k-means on security standards uncovers hidden patterns, expanding the risk assessment scope. Additionally, machine learning enhances risk discovery and classification, while the introduction of heat maps and a thesaurus for risk descriptions innovates visualization and semantic accuracy. The development of an ontological model offers an effective, systematic method for categorizing risks. Together, these advancements significantly improve the management of information security risks, marking a major step forward in the field and setting the stage for future developments in adapting to an evolving technological environment. The model developed in this paper can be applied across various industry sectors to enhance information security risk assessment and management. The healthcare sector deals with highly sensitive data, including personal health information that must be protected from breaches and unauthorized access. The model can help healthcare organizations assess and manage risks related to data privacy and security. Banks and financial institutions are prime targets for cyber attacks due to the valuable financial and personal data they hold. The model can assist in identifying and prioritizing risks associated with online transactions, data storage, and third-party services, helping these institutions strengthen their cybersecurity measures and comply with financial regulations. As manufacturing becomes more digitized with the adoption of IoT devices and automated systems, it faces new vulnerabilities. The model can be used to assess risks associated with the integration of these technologies and to develop robust security frameworks to protect against potential cyber threats. Educational institutions store vast amounts of personal information about students and staff. The model can guide these institutions in assessing the risks associated with data storage and access, especially as e-learning platforms become more prevalent.

**Author Contributions:** Conceptualization, A.B. and A.S.; methodology, A.B.; software, A.B.; validation, A.S., G.S. and A.Z.; formal analysis, A.Z.; investigation, A.B.; resources, G.S.; data curation, A.S.; writing—original draft preparation, A.B.; writing—review and editing, A.B.; visualization, G.S.; supervision, A.S.; project administration, A.Z.; funding acquisition, A.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grant No. AP19174390).

**Data Availability Statement:** The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding authors.

**Conflicts of Interest:** The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References


1. Abdymanapov, S.A.; Muratbekov, M.; Altynbek, S.; Barlybayev, A. Fuzzy Expert System of Information Security Risk Assessment on the Example of Analysis Learning Management Systems. *IEEE Access* **2021**, *9*, 156556–156565. [CrossRef]
2. de Azambuja AJ, G.; Plesker, C.; Schützer, K.; Anderl, R.; Schleich, B.; Almeida, V.R. Artificial Intelligence-Based Cyber Security in the Context of Industry 4.0—A Survey. *Electronics* **2023**, *12*, 1920. [CrossRef]
3. Koolen, C.; Wuyts, K.; Joosen, W.; Valcke, P. From insight to compliance: Appropriate technical and organisational security measures through the lens of cybersecurity maturity models. *Comput. Law Secur. Rev.* **2024**, *52*, 105914. [CrossRef]
4. Santos-Olmo, A.; Sánchez, L.E.; Rosado, D.G.; Serrano, M.A.; Blanco, C.; Mouratidis, H.; Fernández-Medina, E. Towards an integrated risk analysis security framework according to a systematic analysis of existing proposals. *Front. Comput. Sci.* **2024**, *18*, 183808. [CrossRef]
5. AL-Dosari, K.; Fetais, N. Risk-Management Framework and Information-Security Systems for Small and Medium Enterprises (SMEs): A Meta-Analysis Approach. *Electronics* **2023**, *12*, 3629. [CrossRef]
6. Shaikh, F.A.; Siponen, M. Information security risk assessments following cybersecurity breaches: The mediating role of top management attention to cybersecurity. *Comput. Secur.* **2023**, *124*, 102974. [CrossRef]
7. Dong, T.; Zhu, S.; Oliveira, M.; Luo, X.R. Making better IS security investment decisions: Discovering the cost of data breach announcements during the COVID-19 pandemic. *Ind. Manag. Data Syst.* **2022**, *123*, 630–652. [CrossRef]
8. Majumdar, A.; Singh, P. Analysis and impact of COVID-19 disclosures: Is IT-services different from others? *Ind. Manag. Data Syst.* **2023**, *123*, 345–366. [CrossRef]
9. Paulose, H.; Sethi, A. A Survey on Human Behavioral Cybersecurity Risk During and Post Pandemic World. In *International Conference on Innovative Computing and Communication*; Springer Nature: Singapore, 2023; pp. 467–481.
10. Palko, D.; Babenko, T.; Bigdan, A.; Kiktev, N.; Hutsol, T.; Kuboń, M.; Hnatiienko, H.; Tabor, S.; Gorbovy, O.; Borusiewicz, A. Cyber Security Risk Modeling in Distributed Information Systems. *Appl. Sci.* **2023**, *13*, 2393. [CrossRef]
11. Kulshrestha, V.; Verma, S.; Krishna, C.R. Hybrid probabilistic triple encryption approach for data security in cloud computing. *Int. J. Adv. Intell. Paradig.* **2022**, *21*, 158–173. [CrossRef]
12. Barraza de la Paz, J.V.; Rodríguez-Picón, L.A.; Morales-Rocha, V.; Torres-Argüelles, S.V. A Systematic Review of Risk Management Methodologies for Complex Organizations in Industry 4.0 and 5.0. *Systems* **2023**, *11*, 218. [CrossRef]
13. Biswas, B.; Mukhopadhyay, A.; Kumar, A.; Delen, D. A hybrid framework using explainable AI (XAI) in cyber-risk management for defence and recovery against phishing attacks. *Decis. Support Syst.* **2023**, *177*, 114102. [CrossRef]
14. Al-Dhaqm, A.; Othman, S.H.; Yafooz, W.M.; Ali, A. Review of Information Security Management Frameworks. In *Kids Cybersecurity Using Computational Intelligence Techniques*; Springer International Publishing: Cham, Switzerland, 2023; pp. 69–80.
15. Kure, H.I.; Islam, S.; Mouratidis, H. An integrated cyber security risk management framework and risk predication for the critical infrastructure protection. *Neural Comput. Appl.* **2022**, *34*, 15241–15271. [CrossRef]
16. Mažeika, D.; Butleris, R. Integrating security requirements engineering into MBSE: Profile and guidelines. *Secur. Commun. Netw.* **2020**, *2020*, 5137625. [CrossRef]
17. Zammani, M.; Razali, R.; Singh, D. Organisational information security management maturity model. *Int. J. Adv. Comput. Sci. Appl.* **2021**, *12*, 668–678. [CrossRef]
18. Blancaflor, E.; Banzon CV, H.; Jackson CJ, J.; Jamena, J.N.; Miraflores, J.; Samala, L.K. Risk assessments of social engineering attacks and set controls in an online education environment. In Proceedings of the 2021 3rd International Conference on Modern Educational Technology, Jakarta, Indonesia, 21–23 May 2021; pp. 69–74.
19. Huang, Y.; Li, Y.J.; Cai, Z. Security and privacy in metaverse: A comprehensive survey. *Big Data Min. Anal.* **2023**, *6*, 234–247. [CrossRef]
20. Kumar, M.; Maple, C.; Chand, S. An efficient and secure identity-based integrity auditing scheme for sensitive data with anti-replacement attack on multi-cloud storage. *J. King Saud Univ. Comput. Inf. Sci.* **2023**, *35*, 101745. [CrossRef]
21. Depuru, S.; Hari, P.; Suhaas, P.; Basha, S.R.; Girish, R.; Raju, P.K. A Machine Learning based Malware Classification Framework. In Proceedings of the 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 23–25 January 2023; pp. 1138–1143.
22. Saxena, D.; Gupta, I.; Gupta, R.; Singh, A.K.; Wen, X. An AI-driven VM threat prediction model for multi-risks analysis-based cloud cybersecurity. *IEEE Trans. Syst. Man Cybern. Syst.* **2023**, *53*, 6815–6827. [CrossRef]
23. Charoo, N.A.; Khan, M.A.; Rahman, Z. Data integrity issues in pharmaceutical industry: Common observations, challenges and mitigations strategies. *Int. J. Pharm.* **2023**, *631*, 122503. [CrossRef]
24. Firoozjaei, M.D.; Mahmoudyar, N.; Baseri, Y.; Ghorbani, A.A. An evaluation framework for industrial control system cyber incidents. *Int. J. Crit. Infrastruct. Prot.* **2022**, *36*, 100487. [CrossRef]
25. Chen, H.; Turel, O.; Yuan, Y. E-waste information security protection motivation: The role of optimism bias. *Inf. Technol. People* **2021**, *35*, 600–620. [CrossRef]

26. Grigoriadis, C.; Berzovitis, A.M.; Stellios, I.; Kotzanikolaou, P. A cybersecurity ontology to support risk information gathering in cyber-physical systems. In *European Symposium on Research in Computer Security*; Springer International Publishing: Cham, Switzerland, 2021; pp. 23–39.
27. Mukhopadhyay, A.; Jain, S. A framework for cyber-risk insurance against ransomware: A mixed-method approach. *Int. J. Inf. Manag.* **2024**, *74*, 102724. [CrossRef]
28. Qureshi, K.N.; O'Keeffe, G.; O'Farrell, S.; Costelloe, G. Cybersecurity Standards Policies for CPS in, I.o.E. In *Cybersecurity Vigilance and Security Engineering of Internet of Everything*; Springer Nature: Cham, Switzerland, 2023; pp. 177–192.
29. Shokry, M.; Awad, A.I.; Abd-Ellah, M.K.; Khalaf, A.A. When Security Risk Assessment Meets Advanced Metering Infrastructure: Identifying the Appropriate Method. *Sustainability* **2023**, *15*, 9812. [CrossRef]
30. Wangen, G.; Snekkenes, E.A. A comparison between business process management and information security management. In Proceedings of the 2014 Federated Conference on Computer Science and Information Systems, Warsaw, Poland, 7–10 September 2014; pp. 901–910.
31. Putra, A.P.; Soewito, B. Integrated Methodology for Information Security Risk Management using ISO 27005:2018 and NIST SP 800-30 for Insurance Sector. *Int. J. Adv. Comput. Sci. Appl.* **2023**, *14*, 625–632. [CrossRef]
32. Perdana, R.S.; Effendy, A.; Garnida, H.; Fidayan, A.; Nazar, F.; Saepudin, D. Security and Risk Assessment of Academic Information System By Using NIST Framework (A Case Study Approach). In Proceedings of the 2022 16th International Conference on Telecommunication Systems, Services, and Applications (TSSA), Lombok, Indonesia, 13–14 October 2022; pp. 1–5.
33. Lohmann, P.A.; Albuquerque, C.; Machado, R. Systematic Literature Review of Threat Modeling Concepts. *ICISSP* **2023**, *1*, 163–173.
34. Abushark, Y.B.; Irshad Khan, A.; Alsolami, F.; Almalawi, A.; Mottahir Alam, M.; Agrawal, A.; Ahmad Khan, R. Cyber security analysis and evaluation for intrusion detection systems. *Comput. Mater. Contin* **2022**, *72*, 1765–1783. [CrossRef]
35. Romanosky, S.; Ablon, L.; Kuehn, A.; Jones, T. Content analysis of cyber insurance policies: How do carriers price cyber risk? *J. Cybersecur.* **2019**, *5*, tyz002. [CrossRef]
36. Turskis, Z.; Goranin, N.; Nurusheva, A.; Boranbayev, S. Information security risk assessment in critical infrastructure: A hybrid MCDM approach. *Informatica* **2019**, *30*, 187–211. [CrossRef]
37. Aldawood, H.; Skinner, G. Reviewing cyber security social engineering training and awareness programs—Pitfalls and ongoing issues. *Future Internet* **2019**, *11*, 73. [CrossRef]
38. Figueira, P.T.; Bravo, C.L.; López, J.L.R. Improving information security risk analysis by including threat-occurrence predictive models. *Comput. Secur.* **2020**, *88*, 101609. [CrossRef]
39. El-Ghamry, A.; Darwish, A.; Hassanien, A.E. An optimized CNN-based intrusion detection system for reducing risks in smart farming. *Internet Things* **2023**, *22*, 100709. [CrossRef]
40. Mashatan, A.; Sangari, M.S.; Dehghani, M. How perceptions of information privacy and security impact consumer trust in crypto-payment: An empirical study. *IEEE Access* **2022**, *10*, 69441–69454. [CrossRef]
41. Fu, R.; Huang, X.; Xue, Y.; Wu, Y.; Tang, Y.; Yue, D. Security assessment for cyber physical distribution power system under intrusion attacks. *IEEE Access* **2018**, *7*, 75615–75628. [CrossRef]
42. Hossain, N.; Das, T.; Islam, T.; Alam Hossain, M. Cyber security risk assessment method for SCADA system. *Inf. Secur. J. A Glob. Perspect.* **2022**, *31*, 499–510. [CrossRef]
43. Mi, J.; Huang, W.; Chen, M.; Zhang, W. A method of entropy weight quantitative risk assessment for the safety and security integration of a typical industrial control system. *IEEE Access* **2021**, *9*, 90919–90932. [CrossRef]
44. Qin, Y.; Peng, Y.; Huang, K.; Zhou, C.; Tian, Y.C. Association analysis-based cybersecurity risk assessment for industrial control systems. *IEEE Syst. J.* **2020**, *15*, 1423–1432. [CrossRef]
45. Pang, Q.; Wang, H.; Xu, Z. Probabilistic linguistic term sets in multi-attribute group decision making. *Inf. Sci.* **2016**, *369*, 128–143. [CrossRef]
46. Bustince, H.; Barrenechea, E.; Pagola, M.; Fernandez, J.; Xu, Z.; Bedregal, B.; De Baets, B. A historical account of types of fuzzy sets and their relationships. *IEEE Trans. Fuzzy Syst.* **2015**, *24*, 179–194. [CrossRef]
47. Zhang, Y.; Xu, Z.; Wang, H.; Liao, H. Consistency-based risk assessment with probabilistic linguistic preference relation. *Appl. Soft Comput.* **2016**, *49*, 817–833. [CrossRef]
48. Sönmez, F.Ö.; Kılıç, B.G. A decision support system for optimal selection of enterprise information security preventative actions. *IEEE Trans. Netw. Serv. Manag.* **2020**, *18*, 3260–3279. [CrossRef]
49. Ren, C.; Xu, Y.; Dai, B.; Zhang, R. An integrated transfer learning method for power system dynamic security assessment of unlearned faults with missing data. *IEEE Trans. Power Syst.* **2021**, *36*, 4856–4859. [CrossRef]
50. Barlybayev, A.; Akimbekova, G. Development Of The Intellectual System For Assessing Information Security Risks. *J. Namib. Stud. Hist. Politics Cult.* **2023**, *35*, 1351–1362.
51. Gao, G.; Li, X.Y.; Zhang, B.; Xiao, W. Information Security Risk Assessment Based on Information Measure and Fuzzy Clustering. *J. Softw.* **2011**, *6*, 2159–2166. [CrossRef]
52. Kolini, F.; Janczewski, L. Clustering and topic modelling: A new approach for analysis of national cyber security strategies. *PACIS* **2017**, *126*, 1–12.

53. Jung, I.-S.; Kim, E.-J.; Kwak, J. Visualization Model for Security Threat Data in Smart Factory based on Heatmap. In *Proceedings of the Korea Information Processing Society Conference*; Korea Information Processing Society: Daejeon, Republic of Korea, 2021; Volume 1, pp. 284–287.
54. Rajpure, A.S.; Bere, S.S. The Survey Paper on Network Security with Its Thesaurus Attacks and feasible Security Technology. *IJRAR-Int. J. Res. Anal. Rev. (IJRAR)* **2019**, *6*, 171–177.
55. Herzog, A.; Shahmehri, N.; Duma, C. An ontology of information security. *Int. J. Inf. Secur. Priv. (IJISP)* **2007**, *1*, 1–23. [CrossRef]