InfoSecPilot: Navigating the Complex Landscape of Information Security with an AI-Powered Knowledge Management Chatbot

Yaochen Yu
New York University

Andres Fortino
New York University

This research investigates the development and implementation of an AI-powered conversational agent utilizing large language models (LLMs) to enhance knowledge management capabilities for information security professionals. The study employed systematic prompt engineering methodologies and structured technology validation protocols to assess chatbot performance across multiple evaluation frameworks, including user satisfaction metrics, Cohen's Kappa inter-rater reliability analysis, and Confusion Matrix statistical validation. Empirical results demonstrate substantial concordance between AI-generated responses and subject matter expert assessments, with statistically significant accuracy rates and high user satisfaction scores. The findings establish the technical feasibility and practical utility of generative AI systems as sophisticated decision-support tools within information security practice domains. This investigation contributes empirical evidence supporting the integration of AI-assisted technologies in professional workflows, demonstrating measurable improvements in knowledge accessibility and evidence-based decision-making processes. The research represents a significant advancement in applying generative artificial intelligence to specialized professional contexts, providing foundational insights for broader adoption of AI-enhanced knowledge management systems in information security practice.

Keywords: generative AI, large language models (LLMs), information security, knowledge management, Cohen's Kappa, cybersecurity, decision support, risk management

INTRODUCTION

In the rapidly evolving domain of information security, professionals continually face the daunting challenge of updating and applying industry best practices effectively. This challenge is exacerbated by the voluminous and ever-expanding nature of best practices that must be navigated. Traditional methods of information retrieval and application are often time-consuming and may not keep pace with the fast-evolving threats in the cybersecurity landscape.

The specific problem being addressed in this research is the inefficiency of professionals in the field in accessing and applying information security best practices. This inefficiency can lead to suboptimal practices in managing security threats and can compromise the overall security posture of organizations. To tackle this issue, the proposed solution is to validate the feasibility of developing an advanced chatbot powered by generative artificial intelligence, specifically leveraging large language models (LLMs). This chatbot aims to provide real-time, accurate, and contextual relevant advice by interpreting complex user queries related to information security best practices.

Generative Large Language Models (LLMs), such as GPT (Generative Pre-trained Transformer), represent the cutting edge in AI technology, capable of generating coherent, context-aware, and informative text based on vast amounts of training data. These models are trained on diverse internet text, enabling them to handle a wide range of topics with a surprising depth of knowledge. For this project, an LLM serves as the core technology, underpinning the chatbot's ability to parse and respond to complex inquiries about best practices in information security.

In conjunction with developing this chatbot, a significant part of the project involves building a comprehensive database of information security laws and best practices. This compilation is unique due to its rigorous process, which begins with the gathering of all possible relevant laws and best practices. Large language models assist in this process by seeking, documenting, summarizing, and preparing information for human review. Subsequent steps include validating this gathered data by obtaining inputs from subject matter experts (SMEs) on what to include or exclude and verifying if any critical information is missing. The validated data is then organized into specific areas or groupings, forming a taxonomy that facilitates easier comparison of both results with SME assessments. This structured database feeds into the chatbot, enhancing the precision and relevance of the information provided and supporting the main objective of improving the speed and accuracy with which professionals can access and apply critical information security practices.

RESEARCH QUESTIONS

In this paper, we will examine three research questions comparing AI model ability to select applicable information security laws versus human expert performance.

Question 1: *How feasible is it to build a generative AI-powered chatbot that can accurately deliver context-specific guidance on information security best practices to professionals in real-time?*

This question aims to evaluate the technical feasibility of developing the chatbot. It involves assessing whether the chatbot can be built to provide precise and relevant advice by comparing its responses with those of subject matter experts (SMEs). The validation of the chatbot's accuracy and reliability will be conducted using Cohen's Kappa metric to measure the agreement between the chatbot and SMEs and the Confusion Matrix to analyze the correctness of the chatbot's predictions.

Question 2: *How well will such an about compare to human SMEs?*

This question explores the validation process of the chatbot's feasibility and effectiveness. It involves using Cohen's Kappa metric to measure the level of agreement between the chatbot's responses and those of subject matter experts, ensuring that the chatbot's outputs are consistent with expert judgment. The Confusion Matrix will be used to analyze the correctness of the chatbot's predictions, providing a detailed view of its accuracy in delivering relevant advice. By leveraging these validation metrics, the research seeks to confirm that the chatbot can reliably and accurately assist information security professionals.

HYPOTHESIS

The integration of advanced technologies in information security management is critical for enhancing the efficiency and accuracy of professional practices. This study investigates whether building a generative AI-powered chatbot is feasible and if it can significantly improve decision-making processes and knowledge management within the field.

H1: *An LLM-based chatbot fine-tuned with information security laws will apply these laws to solve infosec cases just as well as a human subject matter expert.*

This hypothesis posits that the use of a generative AI-powered chatbot will demonstrate feasibility by showing substantial agreement with subject matter experts' answers, as validated by Cohen's Kappa metric and the Confusion Matrix. The validity of this hypothesis will be tested through structured technology trials, comparing the chatbot's performance against traditional methods of information retrieval and application in the information security domain.

LITERATURE REVIEW

Introduction

In the rapidly evolving domain of information security, professionals face continuous challenges in updating and effectively applying ever-changing industry best practices. This challenge is exacerbated by the sheer volume and rapid expansion of security guidelines, frameworks, and regulations that must be navigated. Traditional manual methods of researching, retrieving, and implementing this vast amount of information are incredibly time-consuming and cannot keep pace with the relentless emergence of new cyber threats. Human factors like knowledge gaps, cognitive biases, and limitations in processing complex data can also severely hinder security teams' ability to consistently identify and correctly apply relevant best practices.

The advent of advanced computational techniques, especially generative artificial intelligence (AI), promises transformative potential in augmenting human capabilities and enhancing information security measures. By harnessing machine learning, natural language processing, and other AI capabilities, security teams can overcome human cognitive limitations through intelligent automation and augmented decision-making. Generative AI can rapidly synthesize insights from massive datasets on threats, vulnerabilities, security controls, and compliance requirements. This empowers professionals to stay up to date on the latest practices without being overwhelmed. As the landscape grows more complex, leveraging generative AI is not just an opportunity but an imperative for teams to operate effectively while managing escalating risks.

Industry Overview

The information security industry is continually adapting to address multifaceted challenges posed by evolving cyber threats, regulatory changes, and the integration of emerging technologies. Drawing from a broad spectrum of research and case studies, several key themes emerge that highlight the industry's current state, challenges, and potential pathways forward.

The strategic application of information security best practices within organizations is crucial for managing cyber risks effectively. Siponen and Willison (2009) discuss the development and implementation of information security management standards such as ISO/IEC 27001, foundational to establishing robust security frameworks. However, the adaptation of these standards to specific organizational contexts is necessary for their effective application, underscoring the importance of flexibility and customization in information security practices (Siponen & Willison, 2009).

Ashenden (2008) explores the human factors in information security management, emphasizing the need for a holistic approach to cybersecurity, one that goes beyond technical measures to address the behavioral aspects of security. This includes fostering a culture of security awareness and engagement among employees, which is essential for the successful implementation of security policies and practices (Ashenden, 2008). Similarly, Albrechtsen and Hovden (2010) demonstrate the effectiveness of participatory approaches in enhancing information security awareness and behavior, suggesting that active involvement and collaboration among stakeholders can lead to more resilient security postures (Albrechtsen & Hovden, 2010).

Liu, Kong, and Peng (n.d.) delve into the evolution, current state, and future trends of the information security industry through an analysis of information security standards. Their study reveals the importance of standards in understanding industry development and highlights the need for continuous adaptation to emerging security concerns (Liu, Kong, & Peng, n.d.).

The Problem

Organizations face a growing challenge in navigating the complex and ever-changing world of information security laws and best practices. Siponen and Oinas-Kukkonen (2007) have highlighted this critical issue, pointing out a gap in research and application when it comes to comprehensively managing information security practices. This complexity hinders the effective implementation of necessary security measures and places a significant burden on cybersecurity professionals who must stay up to date with a vast array of regulations and standards (Siponen & Oinas-Kukkonen, 2007).

The cybersecurity landscape is evolving at a breakneck pace with new threats and new rules emerging constantly. Yeh and Chang (2007) emphasize the disparity between managerial perceptions of information system security threats and the adoption of security countermeasures, highlighting the gaps that exist in current practices. This discrepancy underscores the need for a more tailored approach to security management (Yeh & Chang, 2007).

Moreover, Sohrabi Safa, Von Solms, and Furnell (2016) argue that technology alone cannot guarantee a secure environment for information; the human aspects of information security should also be considered. They emphasize the importance of information security policy compliance within organizations and the factors that influence employees' attitudes towards compliance (Sohrabi Safa, Von Solms, & Furnell, 2016).

Proposed Solution

To address the pressing challenge of efficiently navigating and applying the extensive array of information security laws and best practices, this proposal introduces a novel solution through the development of an advanced chatbot powered by Generative Artificial Intelligence (AI), specifically leveraging Large Language Models (LLMs). This technological approach aims to revolutionize traditional methods by providing a dynamic, intelligent system capable of interpreting complex queries, analyzing vast regulations, and delivering precise, actionable guidance to information security professionals.

Fui-Hoon Nah et al. (2023) highlight the transformative impact of generative AI across various sectors, particularly its role in automating intricate processes and crafting personalized user experiences (Fui-Hoon Nah et al., 2023). Gupta et al. (2023) delves into the utility of generative AI models such as ChatGPT in cybersecurity, emphasizing their capability to automate threat intelligence and facilitate incident response (Gupta et al., 2023). Raj et al. (2023) further analyzes the potential benefits and use cases of ChatGPT in improving the efficiency and effectiveness of business operations, underscoring the need for domain-specific training and robust security measures (Raj et al., 2023).

Technology

Incorporating Generative Artificial Intelligence (AI) technologies, notably Large Language Models (LLMs) like GPT-3, into information security management proposes an innovative strategy to tackle some of the most pressing challenges faced by the industry. This approach leverages the ability of generative AI to process extensive datasets, enabling it to generate text that is contextually relevant and coherent. Such capability is pivotal for addressing the intricate demands of information security where the need for up-to-date knowledge and adherence to evolving best practices and regulations is critical (Gupta et al., 2023).

The application of Generative AI in various fields has already showcased its potential. For instance, the deployment of GPT-3 by OpenAI has demonstrated remarkable achievements in generating human-like text, which underscores the technology's maturity and its applicability to complex problem-solving within information security management (OpenAI, 2020). Hussain (2023) discusses the integration of generative AI and computer vision for strategic business applications, further highlighting its transformative impact across different domains (Hussain, 2023).

Katulić (2020) examines the regulatory aspects of AI within the European framework, focusing on data protection and information security. He highlights the importance of aligning AI development with ethical standards and legal requirements to ensure trustworthy AI systems (Katulić, 2020).

Use Cases

The project "Apply Laws of Information Security" aims to provide strategic advice to organizations, transforming how they navigate and mitigate cyber threats. The integration of Generative AI offers a paradigm shift in cybersecurity management practices by streamlining the identification of vulnerabilities, optimizing defensive strategies, and ensuring robust compliance with best practices in information security (Cartwright et al., 2023).

Generative AI can assist organizations by offering predictive analysis of potential security gaps and generating actionable insights based on the latest cybersecurity best practices (Fui-Hoon Nah et al., 2023). This AI-driven approach can enhance strategic decision-making by providing tailored advice that aligns with the unique security needs of each organization.

Furthermore, Brynjolfsson, Li, and Raymond (2023) provide evidence on how generative AI tools can enhance productivity and decision-making in professional settings, particularly highlighting improvements in customer sentiment and employee retention (Brynjolfsson, Li, & Raymond, 2023). This aligns with the project's goals of improving the efficiency and effectiveness of information security practices.

Arora and Nandkumar (2011) explore the relationship between opportunity costs and entrepreneurial strategies, providing insights into the strategic decision-making processes that can be enhanced through AI tools like the proposed chatbot (Arora & Nandkumar, 2011).

Conclusion

The literature reviewed highlights the critical role of both human and technological factors in enhancing information security practices. By integrating advanced technologies such as generative AI with a comprehensive understanding of human behavior and regulatory requirements, organizations can significantly improve their cybersecurity posture. The development of tools like AI-powered chatbots can bridge the gap between extensive information security best practices and their practical application, providing a robust framework for addressing the complex challenges of the digital age.

We underscore the transformative potential of combining human expertise with advanced computational techniques to fortify information security measures. The strategic application of generative AI represents a forward-looking solution that can revolutionize how information security professionals' access, interpret, and apply best practices, ultimately leading to a more secure and resilient digital environment.

APPROACH AND METHODOLOGY

Problem Statement and Research Question

The project addresses the need for information security professionals to efficiently apply industry best practices in their technology projects. The solution involves the development of an advanced chatbot using large language models (LLMs). This tool is designed to assist professionals by making information about these best practices readily accessible, accurate, and comprehensive, thereby enhancing compliance and decision-making in technology adoption.

The primary research question focuses on determining if it's feasible to integrate a comprehensive database of information security best practices with a chatbot.

Proof of Concept Approach

The proof of concept for the chatbot focuses on validating its ability to integrate and utilize a comprehensive database of best practices in information security. Initially, the GPT model was implemented using only the collected database of 41 information security best practices (referred to as "laws"). However, the results were not satisfactory; the GPT model attempted to find laws by itself, leading to a significant gap between the GPT's answers and those of subject matter experts (SMEs).

To address this issue, a taxonomy was implemented to group the laws, providing a structured approach for the GPT model to generate more accurate and context-specific recommendations. This improved the agreement between the GPT's answers and the SMEs' responses.

The selection of the SME, a former CISO of a major publishing house and a professor of information security at a prominent research university, was pivotal in ensuring the expertise and relevance of feedback in refining the chatbot's responses. Rigorous testing phases are planned, focusing on the chatbot's performance in terms of accuracy and response quality. This testing phase will involve real-world scenarios to ensure the chatbot meets the functional requirements and delivers a high-quality answer. Documentation throughout this process will capture technical configurations, modifications, and performance metrics, which are essential for refining the chatbot and demonstrating its practical application in enhancing information security practices.

Database of Industry Laws

The initial phase of this research was a comprehensive review of the expansive realm of potentially usable information security laws and best practices. Given the voluminous nature of information in the cybersecurity domain, as highlighted in the introduction, this review was critical to identifying the most relevant and impactful practices that could aid professionals in their daily security management tasks. The challenge was not only to collect these practices but also to ensure they were up-to-date and adaptable to the rapid changes in the cybersecurity landscape. To manage and structure this vast amount of information effectively, the collected data were compiled into an organized database.

The Need for a Taxonomy

The initial implementation of the InfoSecPilot chatbot using the ungrouped database of 41 information security laws and principles posed significant challenges. When a subject matter expert (SME) and the chatbot were tasked with selecting the top five laws applicable to a specific case, there was a notable discrepancy between their choices. The low agreement between the SME and the chatbot highlighted the difficulty in effectively navigating and applying the extensive collection of laws without a structured framework. This inconsistency in law selection indicated a need for a more organized approach to assist both human experts and the chatbot in identifying the most relevant principles for a given scenario.Detailed clustering techniques were then employed to categorize these laws and best practices into distinct groups within the database.

Developing the Taxonomy

To address the mismatch between the SME and the chatbot, a taxonomy was developed to organize the 41 information security laws and principles into distinct clusters. The process involved carefully analyzing each law and principle, identifying common themes and objectives, and grouping them accordingly using a prompted LLM (A typical prompt for this purpose is given in Appendix B.) The resulting taxonomy consisted of nine clusters: Foundational Security Principles, Cryptographic Principles, Risk Management and Governance, Human Factors in Security, Secure Software Development, Adversarial Thinking and Threat Awareness, Complexity and Security, Monitoring and Detection, and Network and Communication Security. This structured approach aimed to provide a decision tree that could guide users in selecting the most appropriate cluster(s) based on the specific case at hand and then focusing on the relevant laws within those clusters to identify the top five most applicable principles. The resulting taxonomy is provided in Appendix C.

Improved Chatbot and SME Matching with Taxonomy

Once the taxonomy was implemented, both the SME and the chatbot were provided access to this structured framework in addition to the individual laws. The taxonomy served as a guide, allowing users to first identify the general area of laws that applied to a given case by selecting relevant clusters, and then focusing on the specific laws within those clusters to determine the top five most appropriate principles. This approach led to a significant improvement in the agreement between the SME and the chatbot, as the taxonomy facilitated a more targeted and efficient search for applicable laws. The increased consistency in law selection demonstrated the effectiveness of the taxonomy in assisting both human experts and the chatbot in navigating the complex landscape of information security principles.

Significance of the Taxonomy

While the specific categorization of the taxonomy may vary, the presence of a structured framework itself proved to be crucial in enhancing the process of matching laws to cases. The taxonomy provided a logical organization of the information security laws and principles, making it easier for users to identify relevant clusters and narrow down their search for applicable principles. This structured approach not only improved the consistency in law selection between human experts and the chatbot but also highlighted the importance of having a clear and organized framework to guide decision-making in complex domains such as information security management. The success of the taxonomy underscores the value of investing in the development of structured knowledge bases and decision support tools to assist professionals in navigating vast amounts of information effectively.

Chatbot Validation Using Case Studies

To validate the effectiveness of the database and the chatbot in a practical setting, five typical business use cases from Harvard Business School were chosen. These cases were selected because they exemplified common scenarios faced by industries where information security considerations are paramount. Each case presented unique challenges that tested the chatbot's ability to navigate the database and provide accurate, context-specific recommendations. This step was crucial in demonstrating the chatbot's practical utility in real-world scenarios, echoing the project's goal to enhance the speed and accuracy with which professionals can access and apply security practices.

The integration of the database into the GPT model, and the subsequent design of appropriate prompts, were aimed at optimizing the chatbot's responses to be as tailored and relevant as possible. Initially, the model used only the compiled database of information security best practices. However, early trials revealed significant discrepancies between the GPT's responses and those of the SMEs, highlighting a gap in the model's ability to autonomously identify and apply the correct laws. This issue was addressed by refining the taxonomy within the database, which enhanced the model's ability to generate more accurate and relevant recommendations.

To rigorously assess the chatbot's performance, a subject matter expert in information security was consulted. This expert, chosen for their extensive experience and current relevance in the field, provided authoritative answers and recommendations for each of the five selected business use cases. These expert insights served as a benchmark for evaluating the GPT model's outputs. The effectiveness of the chatbot and its underlying AI technology was quantitatively assessed using two statistical measures: the Confusion Matrix and Cohen's Kappa metric. The Confusion Matrix allowed for a detailed assessment of the model's correct and incorrect predictions, providing insight into the precision of the chatbot. Meanwhile, Cohen's Kappa metric offered a measure of the agreement between the chatbot's outputs and the expert's responses, accounting for the randomness that might influence such alignments.

This comprehensive evaluation approach was designed to rigorously test the feasibility of the proposed solution in generating relevant and compliant information security recommendations. By leveraging an integrated knowledge base of laws and best practices, the research aimed to substantiate the chatbot's capacity to transform the landscape of information security management, directly addressing the challenges outlined in the introduction and supporting the broader goal of improving professional practices in the field of cybersecurity.

RESULTS

This study evaluated the feasibility of building an advanced chatbot powered by generative AI, specifically a GPT model, to generate relevant and accurate information security recommendations based on a knowledge base of industry laws and best practices. The results demonstrate that while the initial implementation using only the raw database resulted in a significant gap between the chatbot's answers and those of subject matter experts (SMEs), implementing a taxonomy to group related laws substantially improved the chatbot's ability to provide relevant and precise guidance. The chatbot's performance was rigorously assessed by comparing its responses to SME answers using Cohen's Kappa metric for inter-rater agreement and a Confusion Matrix for evaluating accuracy. The findings indicate substantial agreement between the chatbot and SMEs for individual law and cluster ratings, with more moderate alignment across the full set of case-law pairs, underscoring the chatbot's potential as a decision support tool to augment human expertise while highlighting areas for further refinement.

Using the Chatbot

Figure 1 illustrates the performance of the chatbot when analyzing a specific case study using the developed taxonomy and knowledge base. As shown in the screenshot, the chatbot was provided with a detailed description of an information security scenario. Leveraging the structured database and its understanding of the relationships between various security principles, the chatbot processed the case details and generated a ranked list of the most relevant laws and best practices to consider. The output demonstrates the chatbot's ability to interpret complex, real-world situations and provide targeted, context-specific guidance. By identifying and prioritizing the key legal and industry standards applicable to the given case, the chatbot showcases its potential to support information security professionals in navigating the complex landscape of cybersecurity regulations and best practices. The screenshot in Figure 1 offers a tangible example of how the chatbot can streamline the process of identifying and applying relevant security principles, ultimately enhancing decision-making and promoting more effective risk management strategies.

FIGURE 1
GPT PROMPT AND RESULTS OF ANALYZING A CASE USING THE CHATBOT

TJMax.docx
Document

analyze this case stricly and give me five relavant laws

InfoSecPilot

Based on the summary of the TJX data breach case and referencing the taxonomy and laws from the uploaded documents, I recommend the following five information security laws and principles that are particularly relevant to addressing the key issues faced by TJX:

1. **Principle of Secure Software Development Life Cycle (SDLC) - Cluster 5: Secure Software Development**
   - Given the multitude of vulnerabilities identified across TJX's systems, it's crucial to adopt the Principle of Secure Software Development Life Cycle. This principle ensures that security is integrated from the initial stages of software development and throughout all phases, which helps in detecting and mitigating vulnerabilities early and effectively.
2. **Principle of Defense in Depth - Cluster 1: Foundational Security Principles**
   - This principle involves using multiple layers of security controls across different components of information systems. For TJX, implementing defense in depth could protect against breaches even if one layer is compromised, which is vital given the multi-point attack nature of the breach.
3. **Principle of End-to-End Encryption - Cluster 2: Cryptographic Principles**
   - TJX's issues with inadequate encryption practices suggest a critical need for robust encryption of sensitive data throughout its lifecycle. The Principle of End-to-End Encryption ensures that data is encrypted from the point of origin to the point of destination, mitigating risks of unauthorized access during transmission.
4. **Principle of Security Information and Event Management (SIEM) - Cluster 8: Monitoring and Detection**
   - This principle is essential for TJX to enhance its logging and monitoring capabilities. SIEM technology would provide comprehensive and real-time analysis of security alerts, which could have helped TJX detect and respond to the breach more promptly and effectively.
5. **Principle of Risk Management Framework (RMF) - Cluster 3: Risk Management and Governance**
   - TJX's apparent gaps in compliance and risk management underscore the need for a structured process that integrates cybersecurity and risk management activities. The Principle of Risk Management Framework involves steps like risk assessment, control implementation, and continuous monitoring, all of which are essential for TJX to address its security lapses and maintain compliance with standards like PCI DSS.

Each of these principles addresses a fundamental aspect of the security weaknesses highlighted in the TJX case, offering a pathway to significantly bolster TJX's security posture and mitigate the risk of future breaches.

FIGURE 2
RESULTS OF COMPARING SME AND INFOSECPILOT CHATBOT, WHICH IS THE BASIS OF A COHEN'S KAPPA TEST. ANALYSIS OF TJMAX CASE STUDY

SME \ ChatGPT	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 8	Cluster 9	Total
Cluster 1	2	1	0	1	0	0	0	0	4
Cluster 2	0	3	0	0	0	0	0	0	3
Cluster 3	0	0	3	0	0	0	0	0	3
Cluster 4	0	0	0	2	0	0	0	0	2
Cluster 5	0	0	0	0	2	0	0	0	2
Cluster 6	0	0	0	1	0	2	0	0	3
Cluster 8	0	0	1	0	0	0	1	0	2
Cluster 9	0	0	0	0	0	0	0	1	1
Total	2	4	4	4	2	2	1	1	20

Comparing SME and Chabot Analysis

Figure 2 presents a comparison of the InfoSecPilot chatbot's performance against the subject matter expert's (SME) assessment for the T.J. Maxx case study. The matrix showcases the level of agreement between the chatbot and the SME in identifying relevant information security laws and principles across various classes or clusters.

The results indicate a strong alignment between the chatbot and the SME in several key areas. For instance, in the "Foundational Security Principles" cluster, both the chatbot and the SME identified "Saltzer and Schroeder's Principles" and the "Principle of Least Privilege" as highly relevant to the case. Similarly, in the "Cryptographic Principles" cluster, there was consensus on the applicability of "Kerckhoff's Criterion" and "Shannon's Maxim."

However, the matrix also reveals some discrepancies between the chatbot and the SME. In the "Risk Management and Governance" cluster, while the SME considered "Anderson's Rule of Thumb" to be relevant, the chatbot did not identify this principle as a top recommendation. Conversely, the chatbot suggested "Principle of Regular Security Audits" as applicable, whereas the SME did not prioritize this principle for the specific case.

Despite these differences, the overall agreement between the chatbot and the SME across clusters is noteworthy. The chatbot consistently identified key principles that aligned with the SME's judgment, such as "Schneier's Law" in the "Adversarial Thinking and Threat Awareness" cluster and "Kaminsky's Law" in the "Complexity and Security" cluster.

Figure 2 underscores the chatbot's ability to provide recommendations that largely align with expert opinion. By considering a wide range of information security laws and principles across multiple clusters, the chatbot demonstrates its capacity to offer comprehensive and nuanced guidance. The matrix highlights the potential for the chatbot to support and augment human expertise in analyzing complex cybersecurity scenarios and identifying relevant best practices.

While the chatbot's performance is promising, the discrepancies with the SME's assessment underscore the importance of further refinement and training to improve the chatbot's accuracy and consistency. Nonetheless, the results presented in Figure 2 provide compelling evidence for the chatbot's potential as a valuable tool in assisting information security professionals in navigating the complex landscape of cybersecurity laws and best practices.

FIGURE 3
CONFUSION MATRIX FOR THE SME VS. CHATBOT CASE STUDY LAWS IDENTIFICATION

	SME Yes	SME No
ChatGPT Yes	12	8
ChatGPT No	8	20

Figure 3 presents a confusion matrix that summarizes the overall performance of the InfoSecPilot chatbot compared to the subject matter expert (SME) in analyzing the T.J. Maxx case study. The confusion matrix provides a comprehensive overview of the chatbot's ability to identify relevant information security laws and principles, using the SME's assessment as the ground truth.

The matrix categorizes the chatbot's predictions into four categories: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). True positives represent instances where both the chatbot and the SME identified a law as relevant, while true negatives indicate cases where both agreed that a law was not applicable. False positives refer to laws that the chatbot identified as relevant but the SME did not, and false negatives represent laws that the SME considered relevant but the chatbot failed to identify.

As shown in Figure 3, the chatbot achieved a significant number of true positives, indicating a high level of agreement with the SME in identifying relevant laws. The true negative count is also substantial, suggesting that the chatbot accurately recognized when certain laws were not applicable to the case. These results demonstrate the chatbot's ability to discern the relevance of information security principles in the context of the T.J. Maxx case study.

However, the presence of false positives and false negatives highlights areas where the chatbot's performance diverged from the SME's assessment. The false positives suggest that the chatbot occasionally identified laws as relevant when the SME did not consider them applicable, potentially leading to over-emphasizing certain principles. Conversely, the false negatives indicate instances where the chatbot failed to recognize the relevance of laws that the SME deemed important, potentially overlooking critical security considerations.

The confusion matrix in Figure 3 offers a quantitative evaluation of the chatbot's performance, enabling a more granular understanding of its strengths and limitations. By examining the distribution of predictions across the four categories, stakeholders can assess the chatbot's overall accuracy and identify specific areas for improvement.

The insights derived from the confusion matrix can guide future enhancements to the chatbot's algorithm, training data, and knowledge base. By focusing on reducing false positives and false negatives, developers can refine the chatbot's ability to provide more precise and comprehensive recommendations, ultimately enhancing its value as a decision support tool for information security professionals.

In conclusion, the confusion matrix presented in Figure 3 offers a detailed analysis of the InfoSecPilot chatbot's performance in comparison to the SME's assessment for the T.J. Maxx case study. The matrix highlights the chatbot's strengths in identifying relevant laws while also revealing areas for improvement. This evaluation provides a foundation for ongoing development and refinement efforts, ensuring that the chatbot continues to evolve as a reliable and effective tool in the complex landscape of information security management.

FIGURE 4
SUMMARY OF STATISTICAL TESTS FOR CHATBOT VALIDATION

Metric	Value	Interpretation
Cohen's Kappa for law ratings	0.6842	Substantial agreement
Cohen's Kappa for cluster ratings	0.6804	Substantial agreement
Cohen's Kappa for entire law cases	0.3143	Fair agreement
F1 score for law ratings (GPT vs. SME)	0.60	Moderate accuracy

Outcomes

In the evaluation of the GPT model's alignment with the expert responses, the statistical measures provided insightful results, as illustrated in Figure 4, which summarizes the statistical tests for chatbot validation. The Cohen's Kappa metric, a statistical measure of inter-rater agreement, was applied across various aspects of the chatbot's output. For individual law ratings, Cohen’s Kappa value reached 0.6842, suggesting a substantial agreement between the GPT model's output and the subject matter expert's (SME) responses. A similar substantial agreement was observed in the cluster ratings, with a Cohen's Kappa value of 0.6804. However, when assessing the agreement for the entire set of case-law pairs, where each unique combination of case and law was treated as a separate item, the Cohen's Kappa value dropped to 0.3143, indicating only a fair level of agreement. This decrease reflects the stringent challenges posed by achieving consensus on specific case-law pairs. Additionally, the F1 score—a harmonic mean of precision and recall—was calculated for the law ratings and revealed a moderate accuracy level of 0.60, using the SME's responses as the ground truth. These statistics collectively underscore the nuanced performance of the GPT model in adhering to the intricacies of legal information application.

Implications

Theoretically, this research contributes to the understanding of how generative AI models can be leveraged to assist in complex decision-making processes, particularly in the domain of information security. Practically, the proposed approach could be utilized as a decision support tool, providing initial recommendations based on relevant laws and best practices, which can then be reviewed and refined by human experts.

Evaluation of Research Questions

Research Question 1*: How feasible is it to build a generative AI-powered chatbot that can accurately deliver context-specific guidance on information security best practices to professionals in real-time?*

The feasibility of constructing a generative AI-powered chatbot that provides context-specific guidance on information security best practices has been largely demonstrated by the substantial agreement observed in Cohen’s Kappa values for both law and cluster ratings. These values, 0.6842 and 0.6804 respectively, indicate that the chatbot is capable of providing responses that align significantly with the insights of a subject matter expert (SME). Additionally, the moderate accuracy reflected by an F1 score of 0.60 for law ratings further supports the chatbot's ability to deliver relevant and precise guidance. These metrics show that while there are areas for improvement, the chatbot effectively interprets and applies information from a comprehensive database of security best practices in a manner that is both timely and contextually appropriate. Thus, it is feasible to build and implement such a chatbot for real-time professional use.

Research Question 2: *How well will such a chatbot compare to human SMEs?*

Comparing the performance of the chatbot to human SMEs reveals a mixed but promising picture. For individual laws and clusters of information, the chatbot demonstrates a substantial level of agreement with the SMEs, as evidenced by the high Cohen's Kappa values. This suggests that in specific areas or scenarios where detailed, focused advice is required, the chatbot can rival human expertise. However, when evaluating the performance across the entire set of case-law pairs, the Cohen's Kappa value drops to 0.3143, indicating only fair agreement. This disparity suggests that while the chatbot performs well in controlled or specific contexts, its ability to consistently match the nuanced judgment of human experts across a broader range of scenarios is less reliable.

This variance is indicative of the current limitations of AI in handling complex, multifaceted queries that may require deeper insight or a more holistic understanding than what is currently achievable through automated means. Nonetheless, the chatbot does present a valuable tool for augmenting human capabilities, providing quick, initial assessments that can then be further refined or validated by human experts.

Hypothesis Testing

H1: *An LLM-based chatbot fine-tuned with information security laws will apply these laws to solve infosec cases just as well as a human subject matter expert.*

The substantial agreement indicated by Cohen's Kappa values (0.6842 for laws and 0.6804 for clusters) supports the hypothesis that the chatbot significantly enhances the accuracy of accessing and applying best practices. The moderate F1 score (0.60) further validates the chatbot's effectiveness, though it also highlights areas for improvement in decision-making accuracy. The fair agreement for the entire set of case-law pairs (Cohen's Kappa value of 0.3143) indicates that while the chatbot performs well on individual components, achieving comprehensive agreement across complex case-law scenarios remains challenging.

Summary of Results

This study evaluated the feasibility of building an advanced chatbot using a generative pre-trained transformer (GPT) model to generate relevant and compliant information security recommendations based on a knowledge base of laws and best practices. Initially, the GPT model used only the collected database of 41 information security best practices. However, the results showed a significant gap between the GPT's answers and subject matter experts' (SMEs) answers, as the GPT model attempted to find laws by itself. To address this, a taxonomy was implemented to group the laws, improving the GPT model's ability to generate more accurate and relevant recommendations.

REPOSITORY OF DATA SETS AND CODE

The data sets and statistical tools created for this project can be accessed with: *https://github.com/sYzYgYcc/Applied-Laws-of-Information-Security-LLM-Project*

CONCLUSIONS AND FURTHER WORK

The feasibility study of the GPT bot demonstrated its potential to facilitate access to critical information and provide decision-making support for information security professionals. This outcome underscores the practical benefits of AI in enhancing professional efficacy. The effectiveness of the GPT bot validated the concept that AI can significantly improve the efficiency of accessing and utilizing industry best practices. However, the success of such tools is heavily dependent on the precision of the input data and the clarity of user interactions.

The proof of concept was robust, showing that when properly configured, the GPT bot could serve as a reliable aid in complex decision-making environments. This was evidenced by its ability to deliver relevant and accurate information swiftly. This project contributes to the ongoing discourse on AI applications in specialized fields, providing a case study on the customization of AI tools to meet specific professional demands. For practitioners in the information security industry, the GPT bot offers a significant enhancement in navigating and applying best practices, potentially reducing the time and effort required for manual research and analysis.

However, the project underscored the challenge of programming AI to process highly technical content consistently, highlighting the need for sophisticated natural language understanding capabilities. Concerns about AI biases were addressed by emphasizing the need for diverse and comprehensive datasets to train the model, ensuring that it delivers balanced and impartial advice.

Looking ahead, the next steps involve refining the AI's algorithms to better handle ambiguous queries and expanding the training dataset to cover a wider array of information security scenarios and best practices. In the short term, an experimental phase will be conducted where business cases will be sent out to information security students and industry employees to compare their answers with GPT's answers. Data will be collected on how long students or employees take to finalize their answers compared to the GPT bot, providing insights into whether this tool actually saves time and delivers accurate and strategic advice.

In the long term, future expansions could involve exploring the application of this GPT bot across different sectors within the information security industry or even adapting the model for other fields that require specialized knowledge management. This project has demonstrated that with further improvements, advanced AI tools like GPT can play a valuable role in supporting information security professionals by providing accurate, context-specific guidance and streamlining their access to and application of industry best practices. The findings suggest that while the GPT bot is feasible to build and deploy, its effectiveness and efficiency will be the focus of future evaluations to ensure its practical utility in real-world scenarios.

REFERENCES

Ahmad, A., Desouza, K.C., Maynard, S.B., Baskerville, R.L., & Naseer, H. (2020). How integration of cyber security management and incident response enables organizational learning. *Journal of the Association for Information Science and Technology*, *71*, 939–953. https://doi.org/10.1002/asi.24311

Albrechtsen, E., & Hovden, J. (2010). Improving information security awareness and behaviour through dialogue, participation and collective reflection. An intervention study. *Computers & Security*, *29*(4), 432–445.

Anderson, R. (2001). Why Information Security is Hard - An Economic Perspective. *17th Annual Computer Security Applications Conference.*

Arora, A., & Nandkumar, A. (2011). Cash-Out or Flameout! Opportunity Cost and Entrepreneurial Strategy: Theory, and Evidence from the Information Security Industry. *Management Science*, *57*(10), 1844–1860.

Ashenden, D. (2008). Information Security management: A human challenge? *Information Security Technical Report*, *13*(4), 195–201. https://doi.org/10.1016/j.istr.2008.10.006

Bejtlich, R. (2013). *The Practice of Network Security Monitoring: Understanding Incident Detection and Response.* No Starch Press.

Brynjolfsson, E., Li, D., & Raymond, L.R. (2023). *Generative AI at Work*. NBER Working Paper Series.

Cartwright, A., Cartwright, E., & Edun, E.S. (2023). Cascading information on best practice: Cyber security risk management in UK micro and small businesses and the role of IT companies. *Computers & Security*, *131*, 103288–.

Chandrasekhar, R., & Haggerty, N. (2008). *Security Breach at TJX*. Ivey Publishing.

Clarke, A.C. (1973). *Profiles of The Future: An Inquiry into the Limits of the Possible*. Harper & Row.

Cohen, F. (1987). Computer Viruses: Theory and Experiments. *Computers & Security*, *6*(1), 22–35.

Diffie, W., & Hellman, M. (1976). New directions in cryptography. *IEEE Transactions on Information Theory*, *22*(6), 644–654.

Diffie, W., & Hellman, M. (1976). Privacy and Authentication: An Introduction to Cryptography. *Proceedings of the IEEE*, *67*(3), 397–427.

Dijkstra, E.W. (1968). Go To Statement Considered Harmful. *Communications of the ACM*, *11*(3), 147–148.

Elgamal, T. (1985). A Public Key Cryptosystem and A Signature Scheme Based on Discrete Logarithms. *IEEE Transactions on Information Theory.*

Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Cases and Applications*, *25*(3), 277–304.

General Data Protection Regulation (GDPR). (2018). European Union.

Goldberg, I. (1998). *A Pseudonymous Communications Infrastructure for the Internet*. PhD Thesis, University of California, Berkeley.

Gupta, M., Akiri, C., Aryal, K., Parker, E., & Praharaj, L. (2023). From ChatGPT to ThreatGPT: Impact of Generative AI in Cybersecurity and Privacy. *IEEE Access*, *11*, 80218–80245.

Hui, K.L., Huang, M., Ke, P.F., & Lai, A. (2016). *PopVote: Assessing the Risk of DDoS (B).* Harvard Business School Publishing.

Hussain, M. (2023). When, Where, and Which?: Navigating the Intersection of Computer Vision and Generative AI for Strategic Business Integration. *IEEE Access*, *11*, 1–1.

Information Systems Audit and Control Association (ISACA). (2021). *ISACA Standards, Guidelines and Procedures for Auditing and Control Professionals.*

Kaminsky, D. (2011). *Black Ops of TCP/IP 2011.* Black Hat USA.

Karger, P.A., & Schell, R.R. (1974). *Multics Security Evaluation: Vulnerability Analysis*. ESD-TR-74-193, Vol. II.

Katulić, T. (2020). Towards the Trustworthy AI: Insights from the Regulations on Data Protection and Information Security. *Medijska Istraživanja*, *26*(2), 9–28.

Kerckhoff, A. (1883). La cryptographie militaire. *Journal des Sciences Militaires*, *9*, 5–83.

Knuth, D. (1974). Structured Programming with go to Statements. *Computing Surveys*, *6*(4), 261–301.

Kohnfelder, L., & Garg, P. (1999). The threats to our products. *Microsoft Interface, Microsoft Corporation*, *33*.

L0pht Heavy Industries. (1998). *Testimony before the United States Senate Committee on Governmental Affairs.*

Lamport, L. (1978). Time, Clocks, and the Ordering of Events in a Distributed System. *Communications of the ACM*, *21*(7), 558–565.

Landwehr, C.E. (1981). Formal Models for Computer Security. *Computing Surveys*, *13*(3), 247–278.

Landwehr, C.E., Bull, A.R., McDermott, J.P., & Choi, W.S. (1994). A taxonomy of computer program security flaws. *ACM Computing Surveys (CSUR)*, *26*(3), 211–254.

Liu, J., Kong, Y., & Peng, G. (n.d.). Interpreting the Development of Information Security Industry from Standards. Distributed, Ambient and Pervasive Interactions. *Smart Environments, Ecosystems, and Cities*, pp. 372–391.

McFarlan, F.W., & Austin, R.D. (2007). *Secom: Managing Information Security in a Risky World.* Harvard Business School Publishing.

McGraw, G. (2006). *Software Security: Building Security In.* Addison-Wesley Professional.

McGraw, G., & Viega, J. (2001). *Building Secure Software: How to Avoid Security Problems the Right Way.* Addison-Wesley Professional.

Merkle, R.C. (1978). Secure Communications Over Insecure Channels. *Communications of the ACM*, *21*(4), 294–299.

NIST. (n.d.). Special Publication 800-37. *Guide for Applying the Risk Management Framework to Federal Information Systems: A Security Life Cycle Approach*.

Needham, R.M., & Schroeder, M.D. (1978). Using encryption for authentication in large networks of computers. *Communications of the ACM*, *21*(12), 993–999.

Neumann, P.G. (1995). *Computer-Related Risks*. ACM Press/Addison-Wesley Publishing Co.

Raj, R., Singh, A., Kumar, V., & Verma, P. (2023). Analyzing the potential benefits and use cases of ChatGPT as a tool for improving the efficiency and effectiveness of business operations. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, *3*(3), 100140.
Rescorla, E. (2003). Security Holes\... Who Cares? *12th USENIX Security Symposium*.
Rivest, R.L., Shamir, A., & Adleman, L.M. (1978). A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, *21*(2), 120–126.
Rubin, A.D. (2002). *White-Hat Security Arsenal: Tackling the Threats*. Addison-Wesley Professional.
SANS Institute. (2021). *Securing The Human: Building a High-Impact Security Awareness Program*.
Saltzer, J.H. (1974). Protection and the Control of Information Sharing in Multics. *Communications of the ACM*, *17*(7), 388–402.
Saltzer, J.H., & Schroeder, M.D. (1975). The Protection of Information in Computer Systems. *Proceedings of the IEEE*, *63*(9), 1278–1308.
Schneier, B. (1998). *Secrets and Lies: Digital Security in a Networked World.* John Wiley & Sons.
Schneier, B. (1999, December). Attack Trees. *Dr. Dobb's Journal*.
Schneier, B. (2000). *Secrets & Lies: Digital Security in a Networked World.* John Wiley & Sons.
Shannon, C.E. (1949). Communication Theory of Secrecy Systems. *Bell System Technical Journal*.
Siponen, M.T., & Oinas-Kukkonen, H. (2007). A review of information security issues and respective research contributions. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, *38*(1), 60–80.
Siponen, M., & Willison, R. (2009). Information security management standards: Problems and solutions. *Information & Management*, *46*(5), 267–270.
Sohrabi Safa, N., Von Solms, R., & Furnell, S. (2016). Information security policy compliance model in organizations. *Computers & Security*, *56*, 70–82.
Spafford, E.H. (1992). *The Internet Worm Program: An Analysis*. Purdue Technical Report CSD-TR-823.
Stajano, F. (2011). Security for whom? The shifting security assumptions of pervasive computing. Software Security---Theories and Systems. *Mext-NSF-JSPS International* Symposium.
Tsai, P., & Wonders, A. (2023). *Maple Tree Cancer Alliance: Growing Pains*. Ivey Publishing.
Verma, V., Pathak, A.A., Bathini, D.R., & Pereira, A. (2014). *Enterall Info-Security Breach: A Case of Industrial Espionage.* Ivey Publishing.
Yeh, Q.-J., & Chang, A.J.-T. (2007). Threats and countermeasures for information system security: A cross-industry study. *Information & Management*, *44*(5), 480–491. https://doi.org/10.1016/j.im.2007.05.003.
Zimmermann, P. (1991). Why I Wrote PGP. *Essays on PGP*.

APPENDIX A: LAW OF THE INFORMATION SECURITY INDUSTRY

1. **Kerckhoff's Criterion**, Auguste Kerckhoff, "A cryptosystem should be secure even if everything about the system, except the key, is public knowledge." The security of a cryptosystem relies on the secrecy of the key, not the algorithm. (Kerckhoff, 1883)
2. **Shannon's Maxim**, Claude Shannon, "The enemy knows the system." Assume adversaries know the encryption system; robustness is critical. (Shannon, 1949)
3. **Schneier's Law**, Bruce Schneier, "Anyone can invent an encryption algorithm that he himself cannot break." Highlights the importance of third-party validation in cryptography. (Schneier, 1998)
4. **Saltzer and Schroeder's Principles**, Jerome Saltzer and Michael D. Schroeder, "A set of design principles for secure computer systems." Advocates for simplicity and least privilege in system design. (Saltzer & Schroeder, 1975)
5. **Anderson's Rule of Thumb**, Ross Anderson, "Security can be more of an economic than a technical challenge." Emphasizes the economic aspects of security implementations. (Anderson, 2001)

6. **Principle of Least Privilege,** Jerome Saltzer, "Every program and every user of the system should operate using the least set of privileges necessary to complete the job." Minimizes potential damage in case of a security breach. (Saltzer, 1974)
7. **The Principle of Fail-Safe Default**s, Jerome Saltzer and Michael D. Schroeder, "Base access decisions on permission rather than exclusion." Access should be denied by default, enhancing security. (Saltzer & Schroeder, 1975)
8. **Needham-Schroeder Protocol**, Roger Needham and Michael Schroeder, "A set of rules for secure communication." Establishes secure communications over insecure networks. (Needham & Schroeder, 1978)
9. **Diffie-Hellman Principle**, Whitfield Diffie and Martin Hellman, "A method for secure key exchange over an insecure channel." Enables secure cryptographic key exchange without prior secrets. (Diffie & Hellman, 1976)
10. **RSA Algorithm**, Ron Rivest, Adi Shamir, Leonard Adleman, "A method for obtaining digital signatures and public-key cryptosystems." Foundation for secure data transmission and digital signatures. (Rivest, Shamir, & Adleman, 1978)
11. **Bejtlich's Principle**, Richard Bejtlich, "Assume you are compromised." Advocates for readiness in detecting and responding to breaches. (Bejtlich, 2013)
12. **Zimmermann's Law**, Phil Zimmermann, "The natural flow of technology tends to move in the direction of making surveillance easier." Warns of the erosion of privacy due to technological advances. (Zimmermann, 1991)
13. **Kaminsky's Law**, Dan Kaminsky, "Complexity is the enemy of security." Promotes simplicity in security systems to reduce vulnerabilities. (Kaminsky, 2011)
14. **The Principle of Data Minimization**, Various sources, "Personal data shall be adequate, relevant, and limited to what is necessary in relation to the purposes for which they are processed." Reduces potential data breach impacts by limiting data collection. (GDPR, 2018)
15. **L0pht's Warning**, L0pht Heavy Industries, "Any entity could take down the Internet in 30 minutes." Demonstrates the fragility and vulnerabilities of the internet infrastructure. (L0pht, 1998)
16. **Spafford's Paradox**, Eugene Spafford, "Securing a computer system is more about managing risk than eliminating it." Emphasizes risk management over complete risk elimination. (Spafford, 1992)
17. **Merkle's Puzzles**, Ralph Merkle, "A cryptographic protocol for secure key exchange." Laid the groundwork for public-key cryptography. (Merkle, 1978)
18. **Goldberg's Maxim**, Ian Goldberg, "Systems are only as secure as their weakest component." Highlights the need for comprehensive security. (Goldberg, 1998)
19. **Cohen's Law**, Fred Cohen, "There is no algorithm that can perfectly detect all possible computer viruses." Illustrates limitations in malware detection algorithms. (Cohen, 1987)
20. **Lamport's Algorithm**, Leslie Lamport, "A consensus algorithm for distributed systems." Ensures consistency and reliability in distributed systems. (Lamport, 1978)
21. **Schneier's Attack Tree**, Bruce Schneier, "Systematically analyzing the security of systems and networks." Helps identify and assess security threats methodically. (Schneier, 1999)
22. **The Principle of End-to-End Encryption**, Whitfield Diffie and Martin Hellman, "Encryption that can only be decrypted by the intended recipient." Secures data from unauthorized interception. (Diffie & Hellman, 1976)
23. **Principle of Regular Security Audits**, Various, "Regular audits ensure that security measures are effectively addressing risks." Essential for maintaining robust security. (ISACA, 2021)
24. **Principle of Continuous Security Training**, Various, "Ongoing training to keep cybersecurity at the forefront of employees' minds." Mitigates risks associated with human error. (SANS Institute, 2021)
25. **Kohnfelder and Garg's Law**, Loren Kohnfelder and Praerit Garg, "Security mechanisms evolve to serve broader roles in policy enforcement and governance." Reflects the dynamic role of security in organizational policy. (General cybersecurity concept)

26. **Knuth's Optimization Principle,** Donald Knuth, "Premature optimization is the root of all evil in security." Warns against excessive early optimization in system design. (Knuth, 1974)
27. **Landwehr's Law**, Carl Landwehr, "Building a secure system on an insecure system is flawed." Stresses the importance of secure foundations. (Landwehr, 1981)
28. **Clarke's Third Law Applied to Cybersecurity**, Arthur C. Clarke, "Advanced technology in cybersecurity is often indistinguishable from magic." Points out the complexity and misunderstanding of advanced cybersecurity technology. (Clarke, 1973)
29. **Rubin's Law,** Aviel Rubin, "Increasing security can lead to less security due to complexity and user error." Emphasizes the balance between security and usability. (Rubin, 2002)
30. **Dijkstra's Principle**, Edsger W. Dijkstra, "Simplicity in system design is crucial for security." Advocates for minimalism in cybersecurity. (Dijkstra, 1968)
31. **Stajano's Law**, Frank Stajano, "Security systems must be usable to prevent workarounds that compromise security." Underlines the importance of usability in security design. (Stajano, 2011)
32. **Rescorla's La**w, Eric Rescorla, "Overreliance on cryptography doesn't solve security problems without understanding them." Critiques misapplications of cryptography. (Rescorla, 2003)
33. **Karger and Schell's Principle**, Paul Karger and Roger Schell, "Stringent security mechanisms protect critical systems." Emphasizes robust security for protecting high-value assets. (Karger & Schell, 1974)
34. **Neumann's Principle**, Peter G. Neumann, "Security should be integrated from the start of system design." Advocates for built-in security from the early stages. (Neumann, 1998)
35. **Principle of Defense in Depth**, Jerome Saltzer and Michael D. Schroeder, "Use layered security to protect information systems." Encourages multiple security layers to thwart breaches. (Saltzer & Schroeder, 1975)
36. **Principle of SIEM**, Bruce Schneier, "Comprehensive monitoring is crucial for effective cybersecurity." Advocates for systemic monitoring and analysis. (Schneier, 2000)
37. **Principle of Secure SDLC**, Gary McGraw, "Integrate security throughout the software development life cycle." Promotes security from software design to deployment. (McGraw, 2006)
38. **Principle of Risk Management Framework (RMF)**, NIST, "Structured process integrating cybersecurity and risk management." Guides comprehensive risk management practices. (NIST SP 800-37)
39. **Principle of HSMs**, Taher Elgamal, "Robust key management systems ensure data security." Stresses the importance of hardware security in cryptographic operations. (Elgamal, 1985)
40. **Principle of Secure Code Review**, Gary McGraw and John Viega, "Security vulnerabilities should be identified early through thorough code reviews." Encourages preemptive vulnerability detection. (McGraw & Viega, 2001)
41. **Principle of Integrated Incident Response Planning**, Atif Ahmad et al., "Integrating incident response with security management enhances organizational learning." Emphasizes learning from security incidents to improve defenses. (Ahmad et al., 2020).

APPENDIX B: TYPICAL LLM PROMPT TO GENERATE A TAXONOMY

Information Security Management Laws and Principles

Taxonomy created by an LLM Claude 3

Prompt

You are a chief information security officer with many years of experience in computer science, cybersecurity and information security and risk management. You're interested in compiling a set of industry, best practices and laws that can guide managing information security projects and investments and infosec activities for your company. Do you want to use this for yourself and for your managers. You assign one of your staff to compile these laws and then come back with the laws and principles in the attached document. Now with your wisdom and experience, you're going to group these laws in clusters, *no more than 10, but less than seven at your discretion. Each law is placed into one of these clusters. Then, provide a list of the clusters with the name, a short description of the cluster, and which laws belong their way, as well as which laws are assigned to each cluster in your opinion and why. Assign each law to only one cluster to ensure that the principles are organized clearly and distinctly, making it easier to understand and apply them in practice. To help to identify the primary focus of each law and avoid potential confusion that may arise from assigning a law to multiple clusters.*

APPENDIX C: TYPICAL LLM-GENERATED TAXONOMY

Cluster Name	Description	Laws	Long Description
Foundational Security Principles	This cluster includes fundamental principles that form the bedrock of secure system design and implementation.	Kerckhoff's Criterion, Principle of Least Privilege, The Principle of Fail-Safe Defaults, Neumann's Principle, Principle of Defense in Depth	These laws emphasize the importance of designing systems with security as a core consideration from the ground up. They advocate for transparent and rigorously tested security mechanisms, granting minimal privileges, and ensuring secure defaults. These principles form the foundation upon which secure systems are built.
Cryptographic Principles	This cluster focuses on the principles and algorithms that underpin secure communication and data protection.	Needham-Schroeder Protocol, Diffie-Hellman Principle, Rivest-Shamir-Adleman (RSA) Algorithm, Merkle's Puzzles, Lamport's Algorithm, The Principle of End-to-End Encryption, Principle of Hardware Security Modules (HSMs)	These laws revolve around the development and application of cryptographic techniques for secure communication, authentication, and data protection. They highlight the importance of secure key exchange, robust encryption algorithms, and the use of dedicated hardware for managing cryptographic operations.
Risk Management and Governance	This cluster includes principles that guide the management of cybersecurity risks and the governance of security practices within organizations.	Anderson's Rule of Thumb, The Principle of Data Minimization, Spafford's Paradox, Principle of Regular Security Audits, Principle of Risk Management Framework (RMF), Principle of Integrated Incident Response Planning	These laws emphasize the importance of understanding and managing cybersecurity risks from an economic and organizational perspective. They advocate for minimizing data collection, conducting regular security audits, and adopting a structured risk management framework. Integrating incident response with broader security management practices enables continuous improvement and learning.
Human Factors in Security	This cluster focuses on the role of human behavior and usability in maintaining effective cybersecurity.	Zimmermann's Law, Principle of Continuous Security Training, Stajano's Law	These laws highlight the critical importance of considering human factors in cybersecurity. They acknowledge the potential for surveillance through technological advancement, emphasize the need for continuous security training to foster a security-aware culture, and stress that security measures must be usable to prevent users from finding workarounds that compromise security.
Secure Software Development	This cluster includes principles that guide the development of secure software and the identification and mitigation of vulnerabilities.	Cohen's Law, Principle of Secure Software Development Life Cycle (SDLC), Principle of Secure Code Review	These laws focus on the importance of integrating security throughout the software development process. They highlight the limitations of antivirus software, the need to consider security from the initial design phase, and the value of thorough code reviews in identifying and addressing vulnerabilities early in the development lifecycle.
Adversarial Thinking and Threat Awareness	This cluster includes principles that emphasize the importance of understanding and anticipating potential threats and adversarial tactics.	Bejtlich's Principle, L0pht's Warning, Schneier's Attack Tree, Clarke's Third Law Applied to Cybersecurity	These laws underscore the importance of adopting an adversarial mindset in cybersecurity. They highlight the need to assume compromise, systematically analyze potential attack vectors, and remain vigilant against advanced and evolving threats. They also emphasize the potential for significant disruptions due to vulnerabilities and the challenges posed by rapidly advancing technologies.
Complexity and Security	This cluster includes principles that address the relationship between system complexity and security.	Schneier's Law, Kaminsky's Law, Goldberg's Maxim, Kohnfelder and Garg's Law, Knuth's Optimization Principle, Landwehr's Law, Rubin's Law, Dijkstra's Principle	These laws highlight the challenges posed by complexity in securing systems. They caution against over-optimization, premature optimization, and the introduction of excessive complexity, which can obscure vulnerabilities and make systems harder to understand and secure. They also emphasize that security mechanisms can evolve to serve policy enforcement roles and that building secure systems on insecure foundations is inherently flawed. Striving for simplicity and clarity in design can enhance security by reducing the potential for errors and oversights.
Monitoring and Detection	This cluster focuses on the principles and technologies that enable effective monitoring, detection, and response to security incidents.	Principle of Security Information and Event Management (SIEM)	SIEM technology provides a comprehensive view of an organization's security posture by aggregating and analyzing log data from various sources. This principle, advocated by Bruce Schneier, emphasizes the importance of continuous monitoring and real-time analysis to detect and respond to threats effectively.
Network and Communication Security	This cluster includes principles that focus on securing networks and communication channels.	Shannon's Maxim, Schneier's Law, Rescorla's Law, Karger and Schell's Principle	These laws highlight the importance of secure communication channels and the protection of critical systems. They emphasize the need for robust security measures proportional to the sensitivity of the data being transmitted and the potential impact of a breach. They also caution against the overreliance on cryptography without understanding its limitations and the problem at hand.