AI driven Security Operation Center

SWISS SCHOOL OF BUSINESS RESEARCH Ph.D. by Portfolio

Candidate: JACOBS Ivan Professor: Dr. Steve Mallon

Contents

1 Background

STE-InfoSec's Cybersecurity Systems and Cybersecurity Services divisions heavily rely on third-party vendor tools to furnish essential software components within Security Operations Centres (SOCs), notably Security Information and Event Management (SIEM) and Security Orchestration, Automation, and Response (SOAR) systems. Current initiatives to introduce analytical methodologies to enhance strategic capabilities in these units remain rudimentary. In the Cybersecurity Systems sector, the organization

faces significant risk by merely acting as a reseller and integrator of third-party SIEM and SOAR solutions, as pricing becomes the primary competitive factor, potentially vulnerable to challenges from rivals offering more advanced analytics-driven solutions. Conversely, in the Cybersecurity Services realm, the necessity to manage vast volumes of ever-evolving complex enterprise data, coupled with a shortage of expertise and a steep learning curve, leads to challenges such as prolonged response times, false positives, alert fatigue, and scalability issues. Additionally, the proliferation of disparate tools and solutions from various vendors across SOCs in the industry complicates efforts to achieve a cohesive view and streamline data and operational workflows effectively.

Presently, proposed analytical approaches within STE-InfoSec fail to address these critical challenges inherent in adopting such methodologies in SOC operations and development. Firstly, a major hindrance lies in the scarcity of labeled data, a prevalent gap not only within the cybersecurity domain but also within the organization itself. The limited availability of cybersecurity-specific machine learning datasets and imbalanced data distribution—where normal behavior significantly outweighs malicious instances—poses a formidable obstacle for machine learning techniques.

Attempts to rectify this imbalance often yield incomplete models with high delivery costs due to extensive human intervention required for data labeling and curation. Secondly, the dynamic nature of cybersecurity data, characterized by constant evolution in enterprise network complexity and adaptive adversaries, leads to rapid obsolescence of analytical approaches. The absence of a robust framework for model evolution and optimization in alignment with business-centric metrics results in inadequate adaptation and acceptance by end-users.

Moreover, most proposed approaches fail to address the challenge posed by the overwhelming tooling landscape, often adding to the complexity by introducing new tools requiring specialized Data Science knowledge for interpretation. Finally, few solutions provide a holistic overview of process optimization encompassing monitoring, threat detection, and operations management. Instead, disjointed analytical interpretations typically lead to increased overhead, false positives, and delayed response times.

To achieve its long-term business objectives, STE-InfoSec must transition towards data-driven, AI solutions that offer comprehensive insights, streamline tool usage, minimize tool proliferation, evolve iteratively based on operational and performance feedback, and optimize human efforts effectively.

Ivan Jacobs, recently appointed as the Vice President and Head of Artificial Intelligence Capability at Singapore Technologies Engineering's cybersecurity division, assumes a pivotal role in shaping the organization's AI strategic direction. Tasked with bolstering AI capabilities for cybersecurity, Ivan collaborates closely with Business Units, providing indispensable technical guidance to formulate AI-related product roadmaps in alignment with business growth and strategies. His leadership extends to mentoring team members, overseeing research and development initiatives in AI for cybersecurity, and managing the implementation of AI platforms and processes.

In this capacity, Ivan is instrumental in advising on the technology stack essential for developing, testing, and deploying AI models tailored to cybersecurity solutions. Leveraging his extensive 22-year career, including over a decade as the Lead Artificial Intelligence Expert at the European Commission, Ivan brings a wealth of experience to this strategic endeavor. His prior roles as a Freelance Machine Learning Consultant and Business Intelligence, Machine Learning Consultant for the public sector have further enriched his expertise.

Ivan has defined a comprehensive strategy Section 7.1, for ST-InfoSec that will elevate the organization to a leadership position in AI within the cybersecurity domain. The primary objective is to equip Large Language Models (LLMs) with versatile capabilities that align with strategic priorities addressing critical aspects of cybersecurity. Key focus areas include enabling seamless integration with diverse cybersecurity tools, necessitating adaptability to new tools and rigorous consideration of security implications for robust security measures. Another critical capability involves enhancing the LLM's autonomy in cybersecurity scenario planning, facilitating adaptability to dynamic environments and evolving threats without reliance on labeled data.

By incorporating mechanisms for continuous learning and adaptation based on real-time data, the LLM can effectively assess and respond to complex cybersecurity scenarios. Additionally, emphasis is placed on orchestrating a swarm of generative agents, refining orchestration capabilities for coordinated action during cyber-attacks while detecting and mitigating adversarial attacks on the swarm. The ultimate goal is to empower the swarm to plan, act, and reason in natural language across diverse and dynamic environments.

These capabilities are integrated and productized to achieve impactful outcomes, notably in deployment at STE Security Operations Centers (SOCs). The focus is on reducing response time, minimizing false positives, and enhancing human involvement in SOC tasks towards more complex responsibilities. Furthermore, the offering to STE customers encompasses a comprehensive solution for SOC operations, encompassing threat hunting, response, training, cyber assistance, and Tactics, Techniques, and Procedures (TTP) discovery. Ensuring user-friendly interfaces, regulatory compliance, and establishing mechanisms for continuous updates and improvements based on user feedback and emerging cybersecurity trends are paramount. This integrated approach aligns with overarching strategic priorities, emphasizing scalability, adaptability, and a holistic solution-oriented approach in the rapidly evolving cybersecurity landscape.

2 Decision making in Security Operation Centers (SOCs)

Security Operations Centers (SOCs) play a critical role in safeguarding organizations against cyber threats by efficiently identifying, analyzing, and responding to security incidents. The decision-making process within SOCs involves several stages,as described in Section 7.2.1, each aimed at ensuring a comprehensive and effective response to potential security breaches. While specific processes may vary depending on the organization's size, industry, and the sophistication of their SOC, a generalized framework commonly observed in many SOCs can be outlined.

Additionally, as shown in Section 7.2.2, the collaborative process in SOC problem-solving and decision-making embodies agility, adaptability, and unity of purpose. By leveraging the collective wisdom and expertise of diverse stakeholders, as described in Section 7.2.3, SOCs can effectively navigate the intricacies of modern threat environments, reflecting the dynamic nature of cybersecurity challenges.

3 AI-Human collaborative decision-making for Security Operations Centers (SOCs)

In response to the flaws, described in Section 7.2.7, identified in the decision-making processes Section 7.2.1, within Security Operations Centers (SOCs), a new approach can be proposed to address these challenges and enhance the effectiveness of SOC operations. Drawing upon these insights a structured and AI-Human collaborative decision-making and problem solving framework tailored to the unique requirements of SOCs can be outlined.

In this collaborative intelligence between humans and AI we aim to enhance human efficiency and facilitate smoother execution of valuable tasks. Typically approached from a utilitarian perspective, this collaboration aids decision-making in uncertainty by assessing decision probabilities and evaluating the utility of potential outcomes. Decision theory guides this process, wherein rational decisions are made by maximizing expected utility, combining probability and utility theories. Kim (2023)

Kim (2023) outlines an ethical decision-making that diverges from conventional approaches as it must consider qualitative factors that resist quantification. While AI excels in calculating Maximum Expected Utility (MEU) and conducting arithmetic operations, human involvement is crucial in discerning ethical standards beyond MEU and determining their applicability in various contexts.

Kim (2023) argues that in ethical decision-making, humans play a vital role in assessing and judging standards beyond MEU, leveraging deductive judgment to navigate qualitative complexities. While AI is proficient in quantitative measurements and computations, it is humans who must deliberate on multifaceted ethical considerations and determine appropriate courses of action.

Thus, while AI contributes to decision-making efficiency, human involvement remains indispensable in interpreting ethical standards and ensuring decisions align with broader moral imperatives. Additionally, it is imperative to develop methods for comprehensively evaluating performance quantification of algorithmic interventions within a broader organizational framework Sankaran, Palomino, Knahl, *et al.* (2022); Kim (2023) .

3.1 AI-Driven Decision Making and Problem Solving Processes

3.1.1 Task Deconstruction and Return on Improved Performance

The evolving landscape of work in cybersecurity demands leaders to grasp and manage a spectrum of labor sources, encompassing not only traditional employees but also gig workers, outsourcers, and smart automation solutions Jesuthasan (2019). To thrive in this dynamic environment, ST-InfoSec must optimize their cybersecurity work ecosystem while ensuring alignment with its overarching purpose and mission.

Deconstructing the Cyber Security Analyst job involves identifying its key elements rather than viewing it as a monolithic role Jesuthasan & Boudreau (2018). This approach reveals optimization patterns that are often obscured within traditional job descriptions.

In the realm of cybersecurity, defining the relationship between work performance and the value it creates is crucial. This relationship, commonly referred to as the "return on improved performance" (ROIP), allows us to understand the effectiveness of various tasks within the Cyber Security Analyst role. Not all tasks yield the same level of value, and understanding the trade-offs between them is essential for optimizing the overall efficiency and effectiveness of cybersecurity operations.

By understanding the ROIP for each task within the Cyber Security Analyst role, we can make informed decisions about resource allocation, prioritize initiatives effectively, and optimize cybersecurity operations to mitigate risks efficiently. This approach enables us to achieve the best possible balance between security effectiveness, operational efficiency, and resource utilization.

Deconstructing the job of a Cyber Security Analyst involves breaking it down into its key elements,as shown in Table 1, rather than treating it as a singular role. This approach allows us to uncover optimization patterns that may be hidden within traditional job descriptions. While the overarching title of "Cyber Security Analyst" may remain intact, the specific tasks and responsibilities within the role can evolve over time. By deconstructing the Cyber Security Analyst job, we identify various essential elements:

- **Repetitive vs. Variable:**
  - Alert Generation, Triage, and Analysis may involve repetitive tasks such as reviewing alerts and conducting investigations. However, the nature of alerts and incidents can vary, requiring analysts to adapt their approach.
  - Incident Response and Continuous Improvement tasks are more variable, as each incident and improvement initiative may present unique challenges and requirements.
- **Independent vs. Interactive:**
  - Alert Generation and Triage tasks can often be performed independently, as analysts review and prioritize alerts based on predefined criteria.
  - Alert Analysis, Incident Response, and Continuous Improvement tasks are more interactive, requiring collaboration with other teams and stakeholders to effectively respond to incidents and drive improvements.
- **Physical vs. Mental:**
  - The tasks within the Cyber Security Analyst role primarily involve mental effort, such as analyzing alerts, coordinating incident response activities, and documenting findings.
  - While there may be some physical aspects, such as operating computer systems and devices, the job's focus is predominantly on cognitive functions and decision-making processes.

Job Name	Performance Standard	Activity Detail	Activity Classification	Possible Job Locations	Time Allocation (in minutes spent)
Cyber Security Analyst	Alert Generation	Generate security alerts using monitoring systems (IDS, SIEM, EDR)	Repetitive, independent, mental	On-site/off-site	10
Cyber Security Analyst	Alert Triage	Assess incoming alerts for severity and prioritize based on predefined criteria	Repetitive, independent, mental	On-site/off-site	15
Cyber Security Analyst	Alert Analysis	Conduct in-depth analysis of high-priority alerts, correlating multiple indicators	Variable, interactive, mental	On-site	30
Cyber Security Analyst	Incident Response	Lead incident response efforts, coordinating with stakeholders to contain, mitigate, and remediate security incidents	Variable, interactive, mental	On-site/off-site	45

Job Name	Performance Standard	Activity Detail	Activity Classification	Possible Job Locations	Time Allocation (in minutes spent)
Cyber Security Analyst	Documentation and Reporting	Document findings, actions taken, and outcomes in incident reports and case management systems	Repetitive, independent, mental	On-site	20
Cyber Security Analyst	Continuous Improvement	Review incident response processes, performance metrics, and lessons learned to identify areas for improvement	Variable, interactive, mental	On-site/off-site	20

Table 1: Activities breakdown and Time Allocation

Defining the ROIP, is crucial in understanding the effectiveness of various tasks within the Cyber Security Analyst role. Not all tasks yield the same level of value, and understanding the trade-offs between them is essential for optimizing the overall efficiency and effectiveness of cybersecurity operations.

By understanding the ROIP for each task within the Cyber Security Analyst role, we can make informed decisions about resource allocation, prioritize initiatives effectively, and optimize their cybersecurity operations to mitigate risks efficiently. This approach enables us to achieve the best possible balance between security effectiveness, operational efficiency, and resource utilization.

Tasks/Work Elements	Return on Improved Performance	CHARACTERISTICS OF THE WORK ELEMENT	Role of Automation	Type of Automation	Repetitive vs. variable	Independent vs. Interactive	Physical vs. Mental
Alert Generation	Reduce mistakes (negative ROIP)	Variable	Augments	Cognitive automation	Variable	Independent	Mental

Tasks/Work Elements	Return on Improved Performance	CHARACTERISTICS OF THE WORK ELEMENT	Role of Automation	Type of Automation	Repetitive vs. variable	Independent vs. Interactive	Physical vs. Mental
Alert Triage	Reduce mistakes (negative ROIP)	Variable	Augments	Cognitive automation	Variable	Independent	Mental
Alert Analysis	Incrementally improve value (incremental ROIP)	Variable	Augments	Cognitive automation	Variable	Interactive	Mental
Incident Response	Exponentially improve value (exponential ROIP)	Variable	Augments	Social robotics	Variable	Interactive	Physical
Documentation & Reporting	Exponentially improve value (exponential ROIP)	Repetitive	Augments	Cognitive automation	Repetitive	Independent	Mental
Continuous Improvement	Incrementally improve value (incremental ROIP)	Variable	Augments	Cognitive automation	Variable	Interactive	Mental

Table 2: Analysis of Activities and Return On Improved Performance (ROIP) and

The results of the analysis,as shown Table 2, suggest that not all tasks yield the same level of value. Tasks such as Incident Response and Continuous Improvement offer exponential returns on improved performance, indicating their critical importance in cybersecurity operations,emphasizing the need for advanced AI solutions capable of automation and augmentation of theses tasks. On the other hand, tasks like Alert Generation and Triage may offer only incremental or negative returns, emphasizing the need for high quality assurance through automation and augmentation.

3.1.2 AI Automation in Cybersecurity Operations

In this section we outline how AI will be utilized to address the flaws identified in Section 7.2.7 by re-defining the decision making process described in Section 3.1.1 as AI driven. AI-driven automation will

revolutionize the traditional processes of alert generation, triage, analysis, incident response, documentation, and continuous improvement within Security Operations Centers (SOCs) as previously described in Section 7.2.9.

3.1.3 Advanced Alert Analysis and Incident Response:

Beyond initial alert generation and triage, AI will revolutionize the way security incidents are detected, analyzed, and responded to. Machine learning algorithms, fueled by vast amounts of data, will empower SOC analysts to swiftly identify and prioritize critical incidents, even amidst a deluge of alerts.

AI-driven incident response capabilities will extend beyond traditional methodologies, incorporating predictive analytics to anticipate potential threats before they fully materialize. By analyzing historical data and correlating it with real-time observations, AI can forecast emerging attack vectors and proactively fortify defenses.

3.1.4 Intelligent Documentation and Reporting:

AI's role in documentation and reporting goes beyond mere automation; it encompasses intelligent data analysis and presentation. Natural Language Processing (NLP) algorithms can extract key insights from incident reports, enabling SOC leaders to glean actionable intelligence and strategic guidance.

Moreover, AI-powered reporting tools can dynamically adapt to changing threat landscapes, providing stakeholders with real-time updates and risk assessments. By synthesizing disparate data sources and streamlining reporting workflows, AI ensures that decision-makers are equipped with timely and relevant information to steer cybersecurity initiatives effectively.

3.1.5 Continuous Improvement and Adaptive Learning:

In the realm of continuous improvement, AI serves as a catalyst for organizational agility and resilience. By analyzing historical performance metrics and identifying areas for enhancement, AI-driven systems facilitate iterative refinement of cybersecurity processes and protocols.

Furthermore, AI's adaptive learning capabilities enable it to evolve in tandem with emerging threats and operational dynamics. Through iterative feedback loops and reinforcement learning algorithms, AI continuously enhances its decision-making prowess, staying one step ahead of adversaries in an ever-changing threat landscape.

3.1.6 SOC Analyst Job Redefinition to AI-driven SOC Manager

The role of the traditional SOC Analyst is being redefined into that of an AI-driven SOC Manager. This transformation involves restructuring the responsibilities to focus on validating the work performed by AI, providing feedback to evolve AI algorithms, and enriching the delivered work by the AI with human intelligence.

The AI-driven SOC Manager will be tasked with overseeing the operation of AI systems responsible for alert generation, triage, analysis, incident response, documentation, and continuous improvement within the Security Operations Center (SOC). Their primary responsibility will involve validating the output generated by AI algorithms, ensuring its accuracy, relevance, and alignment with the organization's security objectives. By leveraging their expertise and domain knowledge, the AI-driven SOC Manager will assess the effectiveness of AI-driven processes and intervene when necessary to correct any discrepancies or errors.

Furthermore, the AI-driven SOC Manager will play a crucial role in providing feedback to AI algorithms to facilitate their continuous improvement and evolution. They will analyze the performance metrics, identify areas for enhancement, and collaborate with AI developers to refine algorithms and optimize decision-making processes. This feedback loop will enable the AI to adapt to emerging threats, evolving attack patterns, and changing organizational requirements, enhancing its effectiveness in mitigating risks and safeguarding the organization's assets.

Additionally, the AI-driven SOC Manager will enrich the output delivered by AI with human intelligence, adding context, insights, and strategic considerations to enhance decision-making capabilities. While AI algorithms excel in processing large volumes of data and identifying patterns, human intuition, creativity, and critical thinking skills are invaluable in interpreting complex situations, assessing potential implications, and formulating strategic responses. By integrating human intelligence with AI-driven automation, the SOC can leverage the strengths of both to achieve optimal outcomes and maximize the value delivered to the organization.

The redefinition of the SOC Analyst role into an AI-driven SOC Manager brings significant value to the organization. By harnessing the power of AI-driven automation, the SOC can streamline operations, improve efficiency, and reduce response times. The AI-driven SOC Manager serves as a bridge between AI systems and human analysts, ensuring the accuracy and effectiveness of AI-driven processes while leveraging human expertise to enrich decision-making and enhance overall cybersecurity posture. With this new role in place, a single human could effectively oversee and manage the entire SOC operations, maximizing resource utilization and optimizing security effectiveness.

In the transition to an AI-driven SOC Manager role, the optimization of tasks such as Alert Analysis and Continuous Improvement yields incremental improvements in the return on improved performance (ROIP). Through AI-optimized alert analysis, the SOC can better discern patterns, anomalies, and relationships among alerts, leading to more accurate threat detection and attribution. Continuous improvement initiatives driven by AI enable the SOC to refine its processes, enhance its response capabilities, and adapt to evolving threat landscapes, resulting in ongoing enhancements to security effectiveness and operational efficiency.

Conversely, tasks like Incident Response and Documentation & Reporting experience exponential improvements in ROIP value with AI integration. AI-driven incident response capabilities enable rapid and effective containment, mitigation, and remediation of security incidents, significantly reducing the potential impact on the organization. Furthermore, AI-enhanced documentation and reporting processes ensure comprehensive and accurate records of incidents, actions taken, and outcomes, facilitating regulatory compliance and post-incident analysis.

By maintaining high standards in AI-optimized Alert Generation and Alert Triage, the SOC can uphold the integrity and reliability of its threat detection and prioritization mechanisms, ensuring that only the most

relevant and high-priority alerts are escalated for further investigation, thereby maximizing the efficiency of response efforts and minimizing false positives. Through the strategic integration of AI across these key SOC functions, organizations can achieve significant improvements in security effectiveness, operational efficiency, and overall ROIP.

3.2 Continuous Improvement within the decision-making processes

An AI-driven approach, bolstered by reinforcement learning, stands to revolutionize how Security Operations Centers (SOCs) leverage data-driven analytics and feedback loops to refine their decision-making processes continuously. Through the integration of reinforcement learning algorithms, SOCs can cultivate a culture of perpetual learning and innovation, iteratively enhancing their protocols to proactively tackle evolving cyber threats.

At the core of this methodology lies AI-driven analytics, empowering SOC managers to extract actionable insights from vast amounts of decision-making data. By analyzing historical incident records, response efficacy metrics, and other performance indicators, AI algorithms identify patterns and trends, offering opportunities for improvement.

For example, AI-powered analytics can dissect past incidents to uncover prevalent attack vectors and tactics employed by malicious actors. Equipped with this knowledge, SOCs can adjust their detection algorithms, response playbooks, and mitigation strategies to bolster defenses against imminent threats. Additionally, AI-driven analytics evaluate the effectiveness of analysts and SOC processes, pinpointing performance bottlenecks and optimization opportunities to streamline operations and reduce response times.

3.2.1 Metrics and Key Performance Indicators (KPIs)

In assessing the effectiveness of the AI-driven decision-making framework, our approach involves establishing clear metrics and key performance indicators (KPIs) that focus on minimizing noise in decisions and optimizing decision processes. One crucial metric is the reduction in noise within decision-making, which reflects the framework's ability to minimize erroneous or irrelevant alerts and streamline the identification of genuine security threats. By quantifying the reduction in noise through the comparison of false positives and false negatives, we can gauge the framework's effectiveness in enhancing the signal-to-noise ratio and improving the efficiency of decision processes.

Another pivotal metric pertains to the time taken from threat detection to mitigation, which highlights the framework's responsiveness in addressing security incidents. By closely monitoring and analyzing the duration between threat detection and effective mitigation actions, we can assess the framework's agility and its impact on reducing the overall response time to security threats. This metric provides valuable insights into the efficiency gains achieved through AI-driven automation and decision-making processes, ultimately enhancing the organization's ability to swiftly respond to evolving cyber threats.

3.3 New AI Paradigm

To achieve the ambitious goals of leveraging AI to revolutionize decision-making processes in Security Operations Centers (SOCs), it's crucial to recognize the limitations of current state-of-the-art algorithms. While these algorithms have made significant progress, they face several challenges that must be addressed for the vision to be fully realized.

A critical limitation lies in the capabilities of existing machine learning models, especially models with transformer architecture,the architecture used by most LLMs. While transformers excel in natural language processing tasks, they struggle with processing diverse inputs beyond text, such as structured data from various security tools and sensors in an SOC environment. This limitation hampers the integration of multiple data sources into a unified decision-making framework.

Moreover, transformer models lack explicit mechanisms for maintaining state and memory over time, hindering their ability to capture temporal dependencies in security events. Without robust memory components, these models struggle to maintain context across stages of decision-making, limiting their ability to provide meaningful insights and recommendations.

Additionally, existing AI models often operate reactively, responding to events as they occur rather than proactively anticipating and planning for future threats. This reactive approach leads to inefficiencies in incident response and resource allocation, as SOC teams struggle to keep pace with emerging threats.

Addressing these challenges requires advancements in AI research and development, exploring new architectures and algorithms that can handle diverse data types, maintain state and memory, and incorporate proactive planning capabilities. Interdisciplinary collaboration between AI researchers, cybersecurity experts, and SOC practitioners is essential to ensure that AI-driven solutions are both technically feasible and effective in real-world SOC environments.

The proposed cognitive architecture, as outlined by LeCun (2022) depicted in Figure 1, offers promising solutions to these challenges. Central to this architecture is the concept of world models, which mimic how humans and animals accumulate knowledge about the world through observation and minimal interactions. These models provide a foundation for predicting outcomes, reasoning, planning, and adapting to new situations efficiently.

We now propose a novel deep learning architecture aimed at addressing the aforementioned challenges in revolutionizing decision-making processes. This architecture integrates advanced techniques for handling diverse data types, maintaining temporal context, and enabling proactive anticipation, fostering efficient and effective AI-driven decision-making in real-world environments. The novel deep learning architecture designed to generalize the computation of multi-modal data and define modal-specific analytical methods. Our proposed Multiplex Deep Learning Models (MGNNs) extends the concept of heterogeneous graphs to achieve Multiplex Graph Neural Networks (MGNNs). In a multiplex graph with both inner and inter-plex edges, various plexes of nodes represent different types of relationships or entities. These connections exist within each plex (inner-plex edges) and between plexes (inter-plex edges). The multiplex graph is denoted as $\mathcal{G} = \mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_K$, where $\mathcal{G}_k$ is the k-th plex of the multiplex graph.

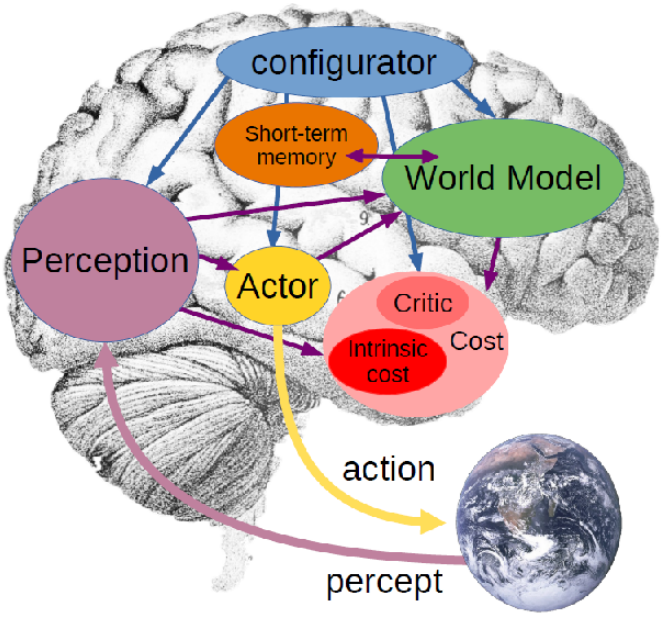

Figure 1

3.3.1 Graph Representation

We introduce the graph representation components for our MGNNs. Let $\mathscr{V}^{(k)}$ be a set of nodes in plex k, $\mathscr{E}^{(k)}$ a set of inner edges in plex k, and $\mathscr{E}^{(k,k')}$ represent the set of inter-plex edges between plexes k and k'. $\mathscr{T}$ is the set of node types, and $\mathscr{R}$ is the set of edge types. The hidden representation of node i of type t in plex k is denoted as $h_i^{(k,t)}$. We use $\mathscr{A}^{(k)}$ as the adjacency matrix for inner edges in plex k and $\mathscr{A}^{(k,k')}$ as the adjacency matrix for inter-plex edges between plexes k and k'. Additionally, $\mathscr{X}^{(k,t)}$ represents the node feature matrix for nodes of type t in plex k.

3.3.2 Inner-Plex Relationships

The inner-plex relationships are computed through the inner-plex adjacency matrix for plex k ($\mathscr{A}^{(k)}$):

$$\mathscr{A}^{(k)} = \begin{bmatrix} a_{11}^{(k)} & a_{12}^{(k)} & \dots & a_{1N}^{(k)} \\ a_{21}^{(k)} & a_{22}^{(k)} & \dots & a_{2N}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1}^{(k)} & a_{N2}^{(k)} & \dots & a_{NN}^{(k)} \end{bmatrix}$$

3.3.3 Inter-Plex Edges

The adjacency matrix for inter-plex edges between plexes k and k' ($\mathscr{A}^{(k,k')}$) captures connections between vertices in different plexes:

$$\mathscr{A}^{(k,k')} = \begin{bmatrix} a_{11}^{(k,k')} & a_{12}^{(k,k')} & \dots & a_{1N}^{(k,k')} \\ a_{21}^{(k,k')} & a_{22}^{(k,k')} & \dots & a_{2N}^{(k,k')} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1}^{(k,k')} & a_{N2}^{(k,k')} & \dots & a_{NN}^{(k,k')} \end{bmatrix}$$

3.3.4 Message Aggregation and Update

Our approach allows the definition of custom heterogeneous message and update functions, empowering the construction of arbitrary Standard Message Passing Graph Neural Networks (MP-GNNs) tailored for heterogeneous graphs from scratch.

3.3.5 Message Aggregation for Inner Edges

For each node type t and its corresponding relations $r \in \mathscr{R}$ within a plex k, we compute the aggregated messages from neighboring nodes, facilitating the incorporation of domain-specific knowledge and task requirements:

$$m_{i \leftarrow j}^{(k,t,r)} = \text{AGGREGATE}\left(\{h_j^{(k,t')}, \forall t', r'\}\right)$$

3.3.6 Message Aggregation for Inter-Plex Edges

Similarly, for each node type t and its corresponding relations $r \in \mathscr{R}$ between plexes k and k', we perform message aggregation from neighboring nodes in different plexes. This operation facilitates the integration of inter-plex dependencies and facilitates information flow across diverse data modalities or relationship types:

$$m_{i \leftarrow j}^{(k,k',t,r)} = \text{AGGREGATE}\left(\{h_j^{(k',t')}, \forall t', r'\}\right)$$

3.3.7 Update Function

Following the message aggregation step, we update the hidden representation of each node based on the aggregated messages. This update function, denoted as UPDATE, allows for the refinement of node representations considering the collected information from neighboring nodes and plexes:

$$h_i^{(k,t)} = \text{UPDATE}\left(h_i^{(k,t)}, \{m_{j \leftarrow i}^{(k,t',r')}\}\right)$$

By allowing the definition of custom message aggregation and update functions, our approach empowers researchers and practitioners to design tailored MP-GNN models that effectively capture the intricate structures and dynamics present in heterogeneous graphs.

3.3.8 Inner and Inter Plex Attention Mechanisms

We introduce inner-plex and inter-plex attention mechanisms in the context of the described multiplex heterogeneous graph neural network:

3.3.8.1 Inner-Plex Attention

The inner-plex attention for node i in plex k can be represented as follows:

$$\text{att}_{ii}^{(k)} = \text{softmax}\left(\text{LeakyReLU}\left(\text{W}_{\text{att}}^{(k)} \cdot \text{emb}_i^{(k)}\right)\right)$$

Here, $\text{W}_{\text{att}}^{(k)}$ is a learnable weight matrix for attention computation.

3.3.8.2 Inter-Plex Attention

The inter-plex attention from node i in plex k to node j in plex l can be represented as:

$$\text{att}_{ij}^{(k\to l)} = \text{softmax}\left(\text{LeakyReLU}\left(\text{W}_{\text{att}}^{(k,l)} \cdot [\text{emb}_i^{(k)}, \text{emb}_j^{(l)}]\right)\right)$$

Here,$\text{W}_{\text{att}}^{(k,l)}$ is a learnable weight matrix for inter-plex attention computation, and $[\text{emb}_i^{(k)}, \text{emb}_j^{(l)}]$ denotes the concatenation of embeddings from plex k and l. The final output is generated based on the learned node representations, which can be utilized for various downstream tasks.

Let's consider a vanilla model with K plexes. The final output $\hat{y}_i$ for node i in plex k is computed using a combination of inner-plex and inter-plex attention mechanisms, as well as a final output layer:

$$\begin{aligned}
\text{Inner-Plex Attention:} \quad & \text{att}_{ii}^{(k)} = \text{softmax}\left(\text{LeakyReLU}\left(\text{W}_{\text{att}}^{(k)} \cdot \text{emb}_i^{(k)}\right)\right) \\
\text{Inter-Plex Attention:} \quad & \text{att}_{ij}^{(k\to l)} = \text{softmax}\left(\text{LeakyReLU}\left(\text{W}_{\text{att}}^{(k,l)} \cdot [\text{emb}_i^{(k)}, \text{emb}_j^{(l)}]\right)\right) \\
\text{Message Aggregation:} \quad & m_{i\leftarrow j}^{(k,l)} = \text{att}_{ij}^{(k\to l)} \cdot \text{AGGREGATE}\left(\{h_j^{(l,t)}, \forall t, r\}\right) \\
\text{Update Function:} \quad & h_i^{(k)} = \text{UPDATE}\left(h_i^{(k)}, \{m_{j\leftarrow i}^{(k,l)}\}\right) \\
\text{Output:} \quad & \hat{y}_i^{(k)} = \text{softmax}\left(\text{W}_{\text{out}}^{(k)} \cdot h_i^{(k)}\right)
\end{aligned}$$

Here, $\text{W}_{\text{att}}^{(k)}$, $\text{W}_{\text{att}}^{(k,l)}$, and $\text{W}_{\text{out}}^{(k)}$ are learnable weight matrices for inner-plex attention, inter-plex attention, and the final plex output, respectively.

This model, as depicted in Figure 2, takes into account both inner-plex and inter-plex relationships, computes attention scores, aggregates messages, updates node representations, and finally produces the output $\hat{y}_i^{(k)}$ using a softmax layer for classification or regression tasks.

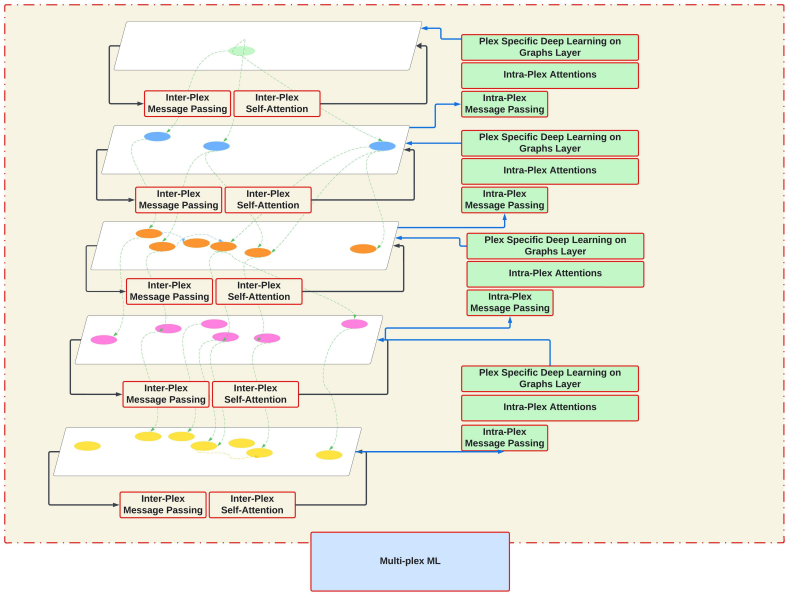

Figure 2

By leveraging on our novel MGNNs we can realize the vision described by LeCun (2022) capable of predicting outcomes, reasoning, planning, and adapting to new situations efficiently. To design an MDM architecture we'll define each module as a plex within the architecture. Let's denote the modules as follows:

- **Perception Module**: Plex P
- **World Model Module**: Plex W
- **Cost Module**: Plex C
- **Short-term Memory Module**: Plex M
- **Actor Module**: Plex A
- **Configurator Module**: Plex Con
- **Critic Module**: (Plex $Crit$)
- **Intrinsic Cost Module** (Plex C)

Now, let's define the functionalities of each module within the multiplex heterogeneous graph neural network architecture:

3.3.9 Perception Module (Plex P)

The perception plex receives signals from sensors and estimates the current state of the world. It represents the state of the world in a hierarchical fashion, extracting relevant information for the task.

3.3.10 World Model Module (Plex W)

The world model plex constitutes the most complex piece of the architecture, serving a dual purpose:

1. Estimating missing information about the state of the world not provided by perception.
2. Predicting plausible future states of the world.

The world model can forecast natural evolutions of the world or anticipate future states resulting from a sequence of actions proposed by the actor module. It may generate multiple plausible world states, characterized by latent variables representing uncertainty about the world state. Essentially, the world model functions as a simulator of relevant aspects of the world, with its relevance contingent upon the specific task. The configurator adjusts the world model to suit the current situation.

Predictions occur within an abstract representation space containing task-relevant information. Ideally, the world model operates across multiple levels of abstraction, facilitating predictions over varying time scales.

A critical challenge lies in enabling the world model to represent multiple possible predictions of the world state. The natural world's unpredictability is compounded by the presence of potentially adversarial intelligent agents or chaotic behavior inanimate objects. Addressing this challenge involves two key questions:

1. How to enable the world model to generate multiple plausible predictions and represent uncertainty in these predictions.
2. How to train the world model effectively.

3.3.11 Cost Module (Plex C)

The cost plex measures the level of discomfort of the agent, computing intrinsic energy and predicting future intrinsic energies. It constitutes the overall objective of the agent, aiming to minimize the average energy over time.

3.3.12 Short-term Memory Module (Plex M)

The short-term memory plex stores relevant information about past, current, and future states of the world, as well as associated intrinsic costs. It interacts with the world model for temporal predictions and with the critic module for training.

3.3.13 Actor Module (Plex A)

The actor plex computes proposals for sequences of actions based on predicted future world states from the world model. It optimizes action sequences to minimize estimated costs and outputs actions to effectors.

3.3.14 Configurator Module (Plex *Con*)

The configurator plex primes other plexes for the task at hand, modulating their parameters and attention circuits. It configures perception, world model, and cost modules to fulfill specific goals and objectives.

3.3.15 Critic Module (Plex *Crit*)

The critic plex predicts an estimate of future intrinsic energies. Its input can be either the current state of the world or possible states predicted by the world model. During training, the critic retrieves past states and subsequent intrinsic costs stored in the associative memory module. It then trains itself to predict the intrinsic costs from the states. The function of the critic module can be dynamically configured by the configurator to direct the system towards a particular sub-goal as part of a bigger task.

3.3.16 Intrinsic Cost Module (Plex *C*)

The intrinsic cost plex is hard-wired, meaning it's immutable and non-trainable. It computes a single scalar value, the intrinsic energy, which quantifies the instantaneous "discomfort" of the agent. This discomfort can represent various states such as pain (high intrinsic energy) or pleasure (low or negative intrinsic energy), hunger, etc. The input to the module is the current state of the world produced by the perception module or potential future states predicted by the world model. The ultimate goal of the agent is to minimize the intrinsic cost over the long run.

Each module in the architecture, as depicted in Figure 3, corresponds to a plex, and interactions between plexes involve inner and inter relationsips,message passing and aggregation, inter and inner attention mechanisms, and updates as described this section. These interactions allow for the integration of information across plexes and facilitate autonomous decision-making and intelligence processes in complex environments.

3.4 Proposed AI-human collaboration paradigm

Leveraging on the new AI-archtitecture we propose in Section 3.3, we envision a new paradigm of AI-human collaboration within Security Operations Centers (SOCs), a human overseer as defined in Section 3.1.6, will orchestrate a swarm of AI agents tasked with handling various functions typically performed by tier one and tier two SOC analysts. These AI agents, operating collectively as a swarm, will possess capabilities for alert generation, triage, incident response, and documentation, mirroring the tasks conducted by human counterparts.

Communication within this framework will be facilitated through natural language, allowing seamless interaction between the human manager and the AI agents. The agents themselves will also engage in natural language exchanges amongst each other to coordinate tasks and share information effectively.

Crucially, the human manager will provide ongoing feedback to the AI agents, guiding their evolution and fine-tuning their performance based on established key performance indicators (KPIs) and metrics.

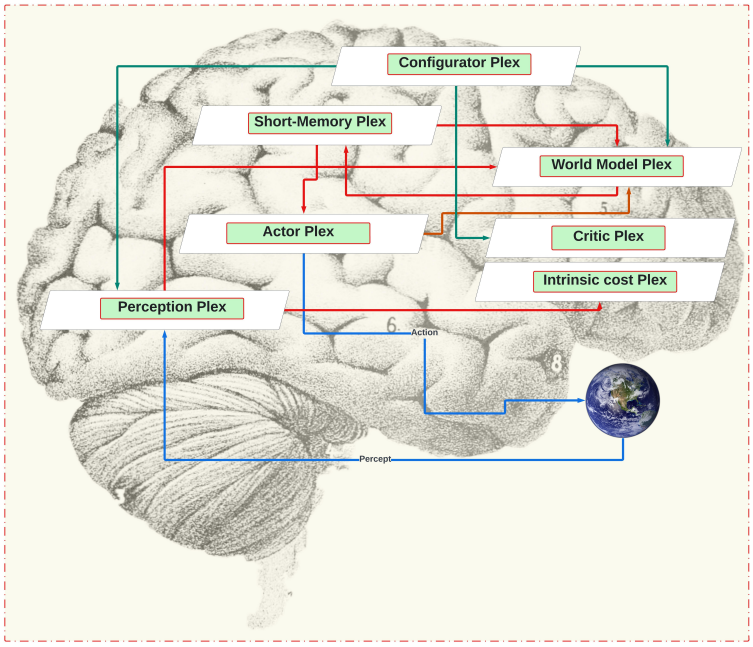

Figure 3

This feedback loop will enable the AI agents to continuously improve and adapt to evolving challenges within the SOC environment.

Furthermore, the AI agents will leverage the same suite of tools utilized by human analysts within the SOC, ensuring compatibility and consistency in operations. This unified approach will maximize efficiency and interoperability between human and AI components, facilitating smooth collaboration and enhancing overall SOC performance.

The proposed AI-human collaboration paradigm redefines traditional decision-making by seamlessly integrating AI capabilities with human expertise. It emphasizes AI as a supportive tool rather, recognizing that while AI excels in data processing and pattern recognition, human judgment remains crucial for interpreting results and making nuanced decisions. This collaborative framework emphasizes synergy between AI-driven insights and human intuition, aiming to leverage the strengths of both entities.

In this approach, AI-driven algorithms act as decision support systems, offering data-driven insights, recommendations, and predictions to inform human decision-makers. These AI agents analyze extensive data, identify patterns, and assess risks, presenting actionable information in a user-friendly format for human interpretation. Human decision-makers critically evaluate AI-generated insights, integrating them with domain knowledge and strategic objectives to make well-informed decisions.

Transparency, explainability, and accountability are pivotal in this collaborative paradigm. AI agents provide transparent explanations in natural language for their recommendations, enabling human decision-makers to understand the rationale behind AI-driven insights and assess their credibility. Human oversight ensures that AI-generated recommendations align with ethical standards and organizational values, mitigating the risks of bias or error.

The AI-human collaboration paradigm signifies a paradigm shift in decision-making, combining the analytical prowess of AI with the cognitive capabilities of human decision-makers. By harnessing the complementary strengths of both entities, this collaborative framework enhances decision quality, fosters innovation, and empowers organizations to navigate complex challenges with resilience and agility.

In the context of Security Operations Centers (SOCs), this collaborative paradigm is particularly transformative. Through advanced AI-driven processes such as alert generation, triage, incident response, and documentation, AI algorithms augment human capabilities to enhance the efficiency and effectiveness of security operations. This fusion of human expertise and AI-driven analytics and automation enables SOCs to detect and respond to security threats with greater speed, accuracy, and adaptability.

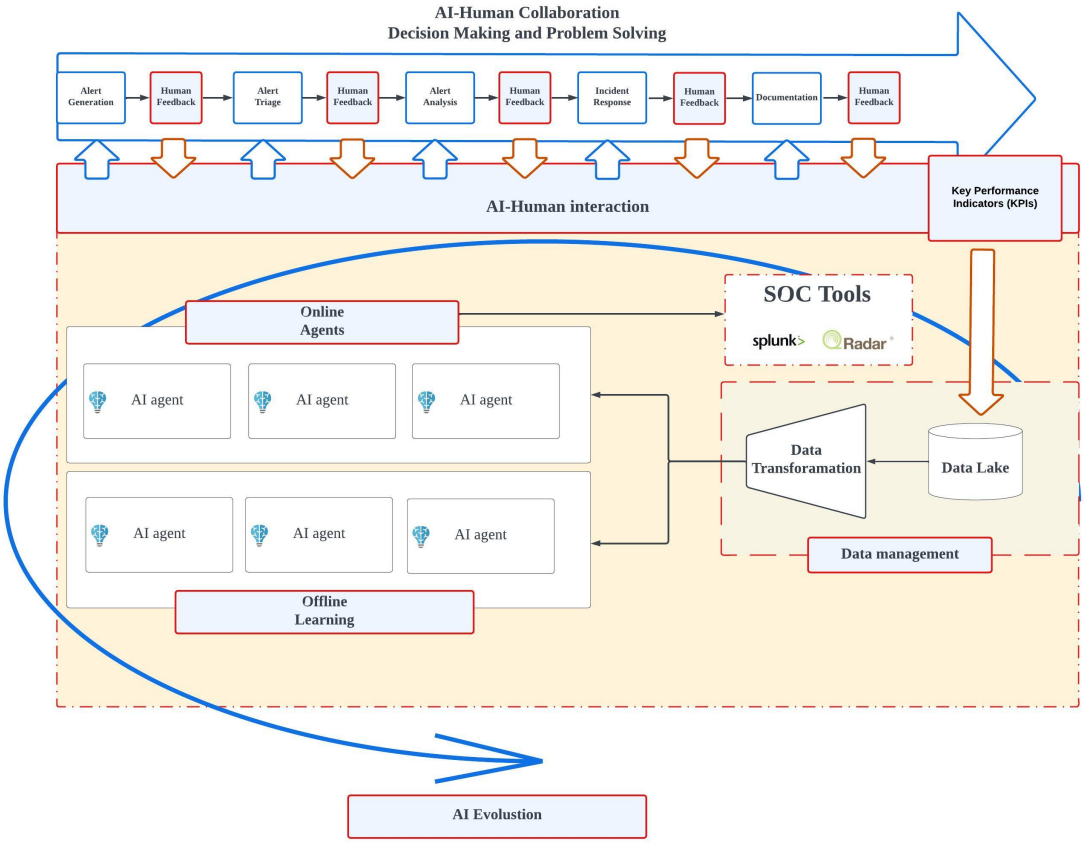

Figure 4

In Section 7.2.12 we show that this paradigm provides a valid solution to the identified flaws. Furthermore,in Section 7.2.7.8 we show that the feasibility assessment, integration challenges, human-machine collaboration, training, and change management aspects of the proposed AI-driven decision-making framework are all aligned with our organization's broader AI strategy, depicted in Section 7.1. This strategy prioritizes the development of infrastructure resources and capabilities to support AI initiatives effectively.

4 Disruptive Innovation

In order for the so far described concept to have growth success we seek to shape it as an disruptive innovation. We assess the disruptive potential by answering three sets of questions as described by Christensen (2013). The first set examines the possibility of a new-market disruption. This occurs if either of the following is true: - Is there a sizable segment of people who lack the means or expertise to access a particular service, leading them to either forgo it entirely or rely on experts? - Does the product or service require customers to visit a centralized, inconvenient location?

If technology can be developed to make something previously accessible only to the affluent or skilled available to a broader population in a more convenient manner, there's potential for a new-market disruption.

The second set of questions delves into the potential for a low-end disruption, which is feasible if: - There are customers at the lower end of the market willing to accept slightly lower performance in exchange for a lower price. - A viable business model can be devised to generate profits despite offering discounted prices to attract these underserved customers.

Low-end disruptions often involve innovations that reduce overhead costs, allowing companies to maintain profitability despite narrower profit margins, along with enhancements in manufacturing or operational processes that increase efficiency.

After an idea passes the new-market or low-end tests, a crucial third question must be affirmed: - Is the innovation disruptive to all major incumbent firms in the industry? If it only benefits one or more established players, the odds favor them, making it difficult for new entrants to succeed.

Applying the principles of disruptive potential assessment Christensen (2013) on the proposed concept we get:

1. **New-market disruption potential**:

    - Is there a sizable segment of people who lack the means or expertise to access SOC services, leading them to forgo it entirely or rely on experts?

      The proposed AI-human collaboration paradigm addresses this point by democratizing access to SOC capabilities through AI-driven automation and natural language communication, potentially reaching a broader population and eliminating the need for SOC analysts physical presence at centralized locations.

      Moreover, Small and Medium Enterprises (SMEs) currently face significant barriers in accessing SOC services due to cost constraints and limited expertise. By leveraging the capabilities of AI-driven automation and natural language communication, the proposed AI-human collaboration paradigm has the potential to unlock this untapped market segment. With SOC analysts no longer bound to physical presence in the SOC, geographical barriers are minimized, enabling ST-InfoSec to outsource this labor to regions where the cost of cybersecurity labor is lower, such as Vietnam, Taiwan, or other emerging markets. This expansion not only broadens the accessibility of SOC services but also taps into new markets previously underserved by traditional SOC models.

2. **Low-end disruption potential**:

   - Are there customers at the lower end of the market willing to accept slightly lower performance in exchange for a lower price?
   - Can a business model be devised to generate profits despite offering discounted prices?

   The AI-driven automation and efficiency enhancements in the proposed paradigm may enable SOC services to be offered at more competitive prices, potentially attracting customers who prioritize cost-effectiveness without compromising on quality.

   Furthermore, the reduction of human involvement in SOC operations, facilitated by AI-driven automation and natural language communication, significantly lowers the overall cost structure of providing SOC services. With fewer human analysts required for day-to-day tasks, the expenses associated with hiring, training, and retaining skilled cybersecurity professionals are notably reduced.

   This reduction in labor costs allows ST-InfoSec to offer SOC services at more competitive prices while still maintaining profitability. Additionally, leveraging AI-driven automation enables ST-InfoSec to streamline workflows, optimize resource allocation, and increase operational efficiency, further contributing to cost savings that can be passed on to customers.

   Consequently, the new business model not only aligns with the cost-conscious preferences of SMEs but also positions ST-InfoSec as a market leader in delivering high-quality SOC services at affordable rates, thereby capturing a larger share of the market and driving sustainable growth.

3. **Disruption to significant incumbent firms**:

   - Is the proposed AI-human collaboration disruptive to all major incumbent firms in the SOC industry?

   If the AI-human collaboration paradigm significantly alters the SOC landscape by offering superior efficiency, effectiveness, and cost-effectiveness compared to traditional methods, it could potentially disrupt established incumbent firms by challenging their existing business models and market dominance.

   In the Security Operations Center (SOC) industry, the traditional approach to AI implementation mirrors a broader trend observed across various sectors, as highlighted in Hammer (1990). This article critiques the prevalent practice of merely integrating innovation within existing paradigms, rather than fundamentally redefining roles and processes.

   Such incremental changes, often characterized by process rationalization and automation, have failed to deliver the transformative improvements demanded by today's rapidly evolving landscape. The incumbent firms approach typically involves augmenting human capabilities with AI tools and technologies, aiming to enhance specific tasks or functions within the SOC framework.

However, in the SOC industry, where the pace of technological change and the importance of efficiency are paramount, incremental improvements may no longer suffice. The proposed AI-human collaboration paradigm represents a departure from this incremental approach, advocating for a radical reengineering of SOC operations.

By leveraging AI-driven automation and natural language communication to orchestrate a swarm of AI agents alongside human overseers, this paradigm offers the potential for quantum leaps in efficiency, effectiveness, and cost-effectiveness.

However, like reengineering efforts in other industries, this transformative shift requires courage and a willingness to break away from outdated processes. Yet, as Hammer (1990) suggest, the potential benefits of reengineering are significant, offering companies the opportunity to shed antiquated practices and thrive in the modern era of innovation and speed.

In conclusion, the proposed AI-human collaboration paradigm for Security Operations Centers (SOCs) demonstrates significant disruptive potential across multiple fronts. By democratizing access to SOC capabilities, particularly for Small and Medium Enterprises (SMEs) and untapped markets, it addresses the barriers of cost and expertise, thus reshaping the SOC landscape.

Moreover, the innovative business model, driven by AI-driven automation, not only enables competitive pricing but also enhances operational efficiency, further solidifying its disruptive impact. This paradigm challenges the traditional incremental approaches of incumbent firms, advocating for a radical reengineering of SOC operations to meet the demands of today's rapidly evolving cybersecurity landscape.

Embracing this transformative shift promises quantum leaps in efficiency, effectiveness, and cost-effectiveness, positioning forward-thinking organizations like ST-InfoSec as leaders in driving sustainable growth and innovation within the SOC industry.

Finally, as outlined by Gilbert & Bower (2002), framing the disruption as a threat during the resource allocation process is crucial to secure sufficient resources. However, once the investment commitment is in place, those involved in venture building must view it as an opportunity for growth. Failing to do so may result in a dangerous lack of flexibility or commitment.

Hence, since ST-InfoSec's senior managers have firmly committed to addressing the disruption, as defined in the accepted AI strategy Section 7.1, the responsibility for commercializing it should be assigned to an independent organizational unit that sees the innovation as a pure opportunity. For ST-InfoSec, securing funding for disruptive growth initiatives marks just the beginning of a continual resource allocation challenge, where the tension between threat and opportunity persists.

Throughout annual budgeting cycles, the disruptive potential may appear marginal, prompting ST-InfoSec to counter with grand future projections to justify present resource allocation. However, this approach is fraught with peril. Firstly, by attempting to quantify potential markets, there's a risk of shoehorning innovation into existing market paradigms, stifling its disruptive potential Kirsner (2021); Christensen (2013). Secondly, if outcomes fail to meet anticipated targets, senior management may perceive the market as limited and subsequently reduce resources. Thus, navigating this dynamic

demands a strategic approach that balances present needs with future potential, avoiding the pitfalls of overpromising and underdelivering Christensen (2013).

5 Consideration of Potential Reactions and Effective Management Strategies

In the following section, we delve deeper into additional considerations essential for successfully implementing the approach outlined in this document, as previously discussed in Section 7.2.7.8.

5.1 Addressing Resistance

Anticipating and addressing resistance to change is essential for successful implementation. Resistance may arise from various quarters, including SOC analysts hesitant about adopting new technologies, managers concerned about job displacement, or IT teams wary of integration challenges. To manage resistance effectively, it's crucial to provide comprehensive training and support programs to equip employees with the necessary skills and knowledge to embrace AI technologies confidently. Additionally, fostering a culture of innovation and continuous learning can help alleviate fears of job displacement by emphasizing opportunities for growth and career advancement in an AI-enabled environment.

5.2 Managing Cultural Shifts

Implementing an AI-driven decision-making framework requires navigating cultural shifts within the organization. Traditional hierarchical structures and decision-making processes may need to adapt to accommodate the collaborative nature of AI-human interaction. Encouraging open communication, collaboration, and experimentation can help foster a culture of trust, adaptability, and innovation conducive to the successful integration of AI technologies. Recognizing and rewarding employees who demonstrate agility, creativity, and adaptability in leveraging AI capabilities can reinforce desired cultural norms and behaviors.

5.3 Proactive Conflict Resolution

Proactively identifying and addressing potential conflicts or disagreements is essential to maintain cohesion and alignment throughout the implementation process. Establishing clear channels for conflict resolution, such as designated escalation paths or mediation mechanisms, can facilitate timely resolution of disputes and prevent issues from escalating. Encouraging constructive dialogue, active listening, and empathy can help foster understanding and consensus among stakeholders with divergent viewpoints.

5.4 Change Management

5.4.1 Stakeholder Engagement

Engaging stakeholders across all levels of the organization is paramount to garner support and alignment with the proposed AI strategy. Transparent communication channels, such as regular town hall meetings, workshops, and feedback sessions, can provide avenues for stakeholders to express their opinions, ask questions, and participate in the decision-making process. Communicating transparently about the rationale, benefits, and implications of the AI-driven decision-making framework will help alleviate concerns and foster buy-in.

5.5 Training and Education

Prioritize comprehensive training and education programs to equip employees with the necessary skills and knowledge to adapt to the AI-driven decision-making framework. Offer hands-on workshops, online courses, and learning resources tailored to different roles and skill levels within the organization.

5.6 Continuous Monitoring and Feedback

Continuous monitoring and feedback mechanisms are critical for assessing the effectiveness of the AI-driven decision-making framework and addressing any emerging issues or concerns promptly. Implementing regular performance evaluations, satisfaction surveys, and post-implementation reviews can provide valuable insights into the framework's impact on SOC operations and employee satisfaction. Soliciting feedback from stakeholders and incorporating their input into ongoing refinements and improvements demonstrates a commitment to responsiveness and continuous improvement, fostering a culture of accountability and engagement.

5.7 Risk Management

5.7.1 Algorithmic Bias

Proactively address algorithmic biases through rigorous testing, validation, and auditing procedures. Ensure diversity in dataset sampling and implement bias detection algorithms to identify and mitigate potential biases in AI models.

5.8 Data Privacy and Security

Strengthen data privacy and security measures to safeguard sensitive information from unauthorized access or breaches. Implement encryption, access controls, and anonymization techniques to protect data integrity and confidentiality.

5.9 Financial Consideration Budget Allocation

Secure adequate budget allocation to support the computational infrastructure required for the AI-driven decision-making framework. Present a detailed business case outlining the expected return on investment (ROI) and the long-term benefits of the AI initiative to justify the budget allocation.

6 Conclusion

The proposed AI-human collaboration paradigm for Security Operations Centers (SOCs) represents a groundbreaking approach poised to revolutionize cybersecurity across various fronts. By democratizing access to SOC capabilities, particularly benefiting Small and Medium Enterprises (SMEs) and untapped markets, it effectively addresses longstanding barriers of cost and expertise, reshaping the landscape of SOC operations. This innovative model, centered on AI-driven automation, not only facilitates competitive pricing but also significantly enhances operational efficiency, marking a departure from traditional incremental approaches and advocating for a radical reengineering of SOC operations to meet the evolving demands of the cybersecurity landscape.

Central to this paradigm shift is the introduction of a new AI architecture, outlined in Section 3.3, which envisions a collaborative framework within SOCs. Here, a human overseer, as defined in Section 3.1.6, orchestrates a swarm of AI agents tasked with handling various functions typically performed by tier one and tier two SOC analysts. These AI agents, operating collectively as a swarm, possess capabilities spanning alert generation, triage, incident response, and documentation, mimicking the tasks traditionally undertaken by human counterparts. This integration of advanced AI capabilities into SOC workflows promises to streamline operations, enhance threat detection and response, and ultimately fortify cybersecurity postures across organizations.

Furthermore, the proposed novel deep learning architecture, as detailed in Section 3.1.6, offers a comprehensive solution to address the complexities inherent in decision-making processes within SOCs. By incorporating advanced techniques for handling diverse data types, maintaining temporal context, and enabling proactive anticipation of threats, this architecture fosters efficient and effective AI-driven decision-making in real-world environments. Leveraging Multiplex Deep Learning Models (MGNNs), which extend the concept of heterogeneous graphs to achieve Multiplex Graph Neural Networks (MGNNs), enables the modeling of intricate relationships and entities within SOC environments. This transformative approach redefines the role of SOC analysts into that of AI-driven SOC Managers, emphasizing validation of AI-generated insights, feedback to evolve algorithms, and augmentation of AI outputs with human intelligence, thereby ensuring a harmonious collaboration between AI and human expertise in SOC operations.

7 Appendices

7.1 Strategy

7.1.1 Vision and Mission

STE-InfoSec aims to be a global leader in cybersecurity innovation by providing comprehensive solutions that empower Security Operations Centers (SOCs) and meet customers' evolving needs. The mission revolves around developing and deploying cutting-edge Large Language Models (LLMs) with seamless integration capabilities, autonomy in cybersecurity scenario planning, and swarm orchestration for effective cyber-attack and defense, focusing on reducing response time, minimizing false positives, and elevating human involvement in SOC tasks.

7.1.2 Strategy

Recognizing the dynamic nature of AI and cybersecurity, STE-InfoSec adopts a flexible strategy emphasizing agility. Investments in advancing AI capabilities aim to ensure the organization's agility in responding to market changes and staying ahead of the curve. Proactively planning for short-term trends, such as legislative developments and advancements in generative AI, enables STE-InfoSec to navigate challenges and capitalize on opportunities swiftly.

7.1.3 Strategic Objectives

STE-InfoSec sets clear objectives to train LLMs for integration with cybersecurity tools, autonomously plan cybersecurity scenarios, and orchestrate a swarm of generative agents. These objectives aim to optimize SOC operations by reducing response time, minimizing false positives, and elevating human involvement. The productization of these objectives focuses on providing comprehensive solutions, user-friendly interfaces, regulatory compliance, and continuous improvement.

7.1.4 Strategic Capability Description

The strategic capabilities revolve around empowering LLMs with versatile capabilities, including integration with multiple APIs, autonomous cybersecurity scenario planning, and swarm orchestration of generative agents. These capabilities align with the organization's goals of scalability, adaptability, and providing holistic solution-oriented approaches. Emphasis is placed on ensuring seamless integration, enhancing reasoning capabilities, refining orchestration, and optimizing deployment for STE SOCs and customers.

7.1.5 Strategic Capabilities List

Integration with Multiple APIs ensures seamless integration with cybersecurity tools, considering security implications. Autonomous Cybersecurity Scenario Planning enhances LLMs' adaptability and reasoning capabilities without relying on labeled data. Swarm Orchestration of Generative Agents refines coordination and mitigates adversarial attacks during cyber-attacks. Deployment at STE SOCs and Offered to STE Customers focuses on optimizing deployment processes, providing clear interfaces, and ensuring scalability. Product Offering for STE Customers includes comprehensive training materials, user-friendly interfaces, and regulatory compliance. Feedback Mechanism and Continuous Improvement gather insights for system enhancement and user-specific needs addressing.

7.1.6 Strategic Priorities

STE-InfoSec articulates its strategy through tangible actions to effectively implement the medium-term plan, emphasizing forward-looking and action-oriented priorities. Holistic Training Framework prioritizes developing a comprehensive training framework for Large Language Models (LLMs), ensuring proficiency in interacting with diverse cybersecurity tools via APIs. Autonomous Cybersecurity Capabilities focus on evolving LLMs to possess autonomous planning, execution, communication, and reasoning, adapting to dynamic cybersecurity scenarios. Swarm Orchestration and Generative Agents Mastery concentrate on designing and training LLMs for orchestrating a swarm of generative agents, ensuring coordinated planning and action.

7.1.7 Operational Impact and Customer-Centric Productization

Operational Impact for SOC Deployment directs efforts toward optimizing system responsiveness to reduce SOC response time and minimize false positives, while elevating human involvement to Tier 3 tasks. Customer-Centric Productization aims to establish an integrated, user-friendly product offering for STE customers, ensuring compliance with cybersecurity regulations and standards and fostering continuous improvement based on user feedback and industry trends.

7.1.8 Strategic Principles

These strategic priorities adhere to simplicity and focus by limiting the number of objectives, channeling efforts and resources with precision. By focusing on mid-term objectives, balancing immediate needs with long-term goals, the organization progresses steadily in a dynamic cybersecurity landscape. Prioritizing autonomous capabilities, swarm orchestration, and customer-centric productization demonstrates a forward-looking approach, positioning the organization at the forefront of emerging technologies.

7.1.9 Decision-Making and Advantages

Making tough decisions to address critical vulnerabilities and optimize SOC responsiveness ensures the organization confronts challenges head-on. Limiting strategic priorities offers advantages in understanding, communication, and retention within the organization, simplifying the strategy for employees and directing attention, effort, and resources towards critical aspects.

7.1.10 Alignment with Future Needs

These strategic priorities embody a balanced approach, aligning with mid-term objectives and preparing for future cybersecurity needs. By focusing on a select few priorities, STE-InfoSec ensures it is well-positioned to navigate the complexities of the cybersecurity landscape effectively.

7.1.11 Conclusion

STE-InfoSec's strategic priorities encompass tangible actions to implement the medium-term plan, emphasizing simplicity, focus, and forward-looking strategies. By prioritizing foundational training, autonomous capabilities, and customer-centric productization, the organization ensures it remains at the forefront of addressing challenges in real-time cybersecurity operations.

7.2 Decision Making

7.2.1 Decision Making in Security Operation Centers (SOCs)

The decision-making process within Security Operations Centers (SOCs) encompasses several stages, starting with alert generation and progressing through triage, analysis, incident response, documentation, and continuous improvement. Alerts are generated by monitoring systems and triaged by Tier 1 analysts based on predefined criteria. Tier 2 analysts conduct in-depth analyses of prioritized alerts, while Tier 3 analysts lead incident response efforts. Comprehensive documentation and reporting are essential throughout the process, facilitating knowledge sharing and accountability. Continuous improvement ensures SOC operations remain effective in addressing evolving threats Onwubiko & Ouazzane (n.d.).

7.2.1.1 Alert Generation

Alert Generation marks the beginning of the decision-making process in SOCs. Security alerts are generated by various monitoring systems, including intrusion detection systems (IDS), security information and event management (SIEM) platforms, endpoint detection and response (EDR) tools, firewalls, and other security devices. These alerts are triggered by predefined rules, anomalies, or suspicious activities detected within the organization's network, systems, or applications.

7.2.1.2 Alert Triage

Alert Triage involves the initial assessment and prioritization of incoming alerts. Level one analysts, also known as Tier 1 analysts, are responsible for this stage. They review incoming alerts, assess their severity, and prioritize them based on predefined criteria such as impact potential, likelihood of exploitation, and relevance to organizational assets. During triage, analysts may perform basic investigations, such as verifying the integrity of network traffic, checking system logs, and conducting preliminary threat assessments.

7.2.1.3 Alert Analysis

Alert Analysis comprises a more in-depth examination of alerts that are deemed to be of higher priority or requiring further investigation. These alerts are escalated to level two analysts, also known as Tier 2 analysts. Tier 2 analysts conduct comprehensive analyses, leveraging additional contextual information and investigative techniques to determine the nature and scope of the security incident. This stage may involve correlating multiple alerts, examining network traffic patterns, conducting memory and disk forensics, and identifying indicators of compromise (IOCs) associated with the incident.

7.2.1.4 Incident Response

Incident Response is initiated once a security incident is confirmed. This stage involves the escalation of the incident to level three analysts, often referred to as Tier 3 analysts or incident responders. Tier 3 analysts lead the incident response efforts, coordinating with various stakeholders, including IT teams, legal counsel, and external vendors, to contain, mitigate, and remediate the incident. Incident response activities may include isolating affected systems, deploying patches or updates, restoring from backups, and conducting post-incident analysis to identify root causes and prevent future occurrences.

7.2.1.5 Documentation and Reporting

Documentation and Reporting are crucial aspects of the decision-making process in SOCs. Throughout the incident response process, analysts document their findings, actions taken, and outcomes in incident reports and case management systems. Comprehensive documentation is essential for maintaining an audit trail, facilitating knowledge sharing, and informing organizational stakeholders about the status of security incidents and the effectiveness of response efforts.

Continuous Improvement is a fundamental principle driving SOC operations. SOC managers and stakeholders regularly review incident response processes, performance metrics, and lessons learned to identify areas for improvement. This may involve refining alerting thresholds, updating response playbooks, enhancing analyst training programs, and investing in new technologies or capabilities to strengthen the organization's security posture.

Overall, the decision-making process in SOCs is characterized by a structured approach to incident detection, analysis, and response, supported by a combination of human expertise, specialized tools, and established procedures. Continuous refinement and adaptation are key principles driving SOC operations to effectively address evolving cyber threats and protect organizational assets.

7.2.2 Collaborative Problem-Solving and Decision-Making

In Security Operations Centers (SOCs), problem-solving and decision-making processes are deeply ingrained within organizational operations, as highlighted by insights from Onwubiko & Ouazzane (n.d.). The collaborative approach is characterized by adaptability and responsiveness, essential qualities for navigating the dynamic cybersecurity landscape. Establishing a cohesive framework that promotes cross-functional collaboration and information sharing is central to this process, fostering open communication and collective responsibility among cybersecurity personnel.

Inc (2017) emphasizes the necessity for increased collaboration between security personnel and operations teams, advocating for a shared SOC model. This model encourages interconnectedness among multiple environments to leverage collective experience and resources in addressing cybersecurity threats. Despite challenges posed by integrating diverse operational technology (OT) architectures, the shared SOC model underscores the significance of synergy in combating cyber threats.

Overall, the collaborative process in SOC problem-solving and decision-making embodies agility, adaptability, and unity of purpose. By leveraging the collective wisdom and expertise of diverse stakeholders, SOCs can effectively navigate the intricacies of modern threat environments, reflecting the dynamic nature of cybersecurity challenges.

7.2.3 Stakeholders in SOC Decision-Making

Various stakeholders contribute to SOC problem-solving and decision-making. Cybersecurity analysts provide real-time monitoring and analysis, while IT operations personnel offer technical expertise in incident response. Executive leadership and management provide strategic guidance, and external partners and vendors augment SOC capabilities with specialized tools and services. This diverse network of stakeholders is united in safeguarding critical assets and infrastructure from cyber threats Inc (2017) ;Dimitrov & Syarova (2019) .

7.2.4 Organizational Culture in a SOC

The collaborative culture within Security Operations Centers (SOCs) is characterized by teamwork, cooperation, and shared goals, fostering open communication and mutual respect among cybersecurity analysts, IT operations personnel, executive leadership, and external partners. This culture promotes collective responsibility for problem-solving and decision-making, facilitating swift identification and mitigation of security incidents Jacq, Boudvin, Brosset, *et al.* (2018); Onwubiko & Ouazzane (n.d.). Additionally, it encourages knowledge sharing, innovation, and adaptability, essential qualities for navigating the complex and dynamic cybersecurity landscape.

Style and people issues, organization culture, and resources and capabilities intertwine to shape the collaborative landscape within SOCs. Factors such as communication styles, conflict resolution mechanisms, and team dynamics influence the effectiveness of collaboration. Fostering a culture of openness and mutual respect is crucial for effective collaboration and knowledge sharing in SOCs.

Organization culture serves as the cornerstone of SOC operations, influencing attitudes, behaviors, and decision-making processes. Whether prioritizing innovation or adherence to established protocols, organizational culture profoundly impacts cybersecurity initiatives. Unique challenges faced by maritime SOCs, as highlighted by Jacq, Boudvin, Brosset, *et al.* (2018), emphasize the importance of prioritization and adaptability in cultivating a conducive organizational culture.

Organization resources and capabilities are pivotal in determining the SOC's ability to address emerging threats. The availability and allocation of resources, including technological infrastructure and human capital, directly impact operational effectiveness. Strategic resource management and investment in training initiatives are essential for enhancing SOC capabilities and resilience against evolving threats.

7.2.5 Style, People Issues, and Organization Culture

In SOCs, fostering a culture of openness and mutual respect is paramount for effective collaboration and knowledge sharing. Organization culture shapes attitudes, behaviors, and decision-making processes, profoundly influencing cybersecurity initiatives. Additionally, addressing challenges such as bandwidth constraints and safety considerations necessitates a culture of prioritization and adaptability, particularly in specialized environments like maritime SOCs Jacq, Boudvin, Brosset, *et al.* (2018); Onwubiko & Ouazzane (n.d.).

7.2.6 Organization Resources and Capabilities

Resources and capabilities significantly influence SOC operations, with staffing levels, processes, and technologies playing crucial roles. Investments in analyst training, incident response playbooks, and technological infrastructure enhance SOC capabilities. However, resource constraints or gaps in capabilities can pose challenges, necessitating strategic resource management and investment in training initiatives Onwubiko & Ouazzane (n.d.); Goel (2010).

7.2.7 Flaws in Decision Making Processes in SOCs

Decision-making in Security Operations Centers (SOCs) faces numerous challenges including alert fatigue, lack of expertise among analysts, and burnout. The reactive nature of SOC operations, compounded by organizational factors like resource constraints and the quality of alerts, further complicates decision-making processes Garofalo (n.d.); Lean Enterprise Institute (2018); Section 7.2.1.

7.2.7.1 Volume of Data and Alert Fatigue

The sheer volume of data inundating SOCs overwhelms analysts, leading to alert fatigue and potentially overlooking critical threats amidst non-critical ones. Additionally, the deluge exacerbates the lack of expertise among analysts, risking misidentification of genuine threats Garofalo (n.d.); Section 7.2.1.

7.2.7.2 Burnout and Reactivity

The relentless nature of SOC work contributes to burnout among analysts, impairing decision-making and increasing the likelihood of errors. The reactive stance limits SOC's ability to prioritize tasks effectively and develop long-term strategic security measures Section 7.2.1.

7.2.7.3 Organizational Factors and Alert Quality

Organizational factors like inadequate resources and unclear escalation procedures hinder effective decision-making. Moreover, the quality of alerts varies significantly, leading to wasted time and eroded trust in the system Garofalo (n.d.); Section 7.2.1.

7.2.7.4 Cognitive Biases and Problem-Solving Methodologies

Cognitive biases like confirmation bias influence analysts' judgment, potentially leading to overlooking critical information. Flaws in problem-solving methodologies within SOCs may result in addressing surface-level symptoms of cyber incidents without delving into underlying systemic issues Garofalo (n.d.); Goel (2010); Lean Enterprise Institute (2018).

7.2.7.5 Lack of Diverse Perspectives and Rapid Technological Evolution

The lack of diverse perspectives and collaborative decision-making processes hinder effective threat detection and response. Additionally, the rapid evolution of technology and the threat landscape pose further challenges to decision-making in SOCs Wujec (n.d.); Goel (2010); Ionescu & Diaconita (2023).

7.2.7.6 Insights from "Noise: A Flaw in Human Judgment"

The book "Noise" highlights cognitive biases and the importance of structured decision-making processes in reducing noise. Technology and data-driven insights, such as artificial intelligence, can assist in mitigating noise and improving decision accuracy in SOCs Kahneman (2021); Davianto (2022).

7.2.7.7 Over-Collaboration and Decision-Making Structures

Over-collaboration in SOC teams can hinder decision-making processes, leading to delays in responding to security incidents. Effective decision-making structures, emphasizing functional expertise and collaborative debate, are crucial for mitigating security threats Shambaugh (2018); Podolny & Hansen (2020).

7.2.7.8 Pitfalls and Recommendations

Flaws in decision-making processes within SOCs, such as over-reliance on management presentations and groupthink, undermine SOC operations' effectiveness. Recommendations include adopting structured decision frameworks and leveraging technology to augment human judgment De Smet (n.d.); Adams (n.d.).

7.2.7.8.1 Feasibility Assessment

ST-InfoSec's AI strategy emphasizes the need for thorough feasibility assessments to ensure the successful implementation of AI-driven solutions. By evaluating our organization's resources, expertise, and readiness, we can make informed decisions about adopting and sustaining the proposed framework.

7.2.7.8.2 Integration Challenges

Addressing integration challenges is a key component of our AI strategy. We recognize the importance of compatibility, data interoperability, and minimizing disruptions during the implementation phase. Developing a detailed integration plan will help us ensure a smooth transition to the new framework.

7.2.7.8.3 Training and Change Management

Training and change management efforts are integral to our AI strategy's success. We are committed to upskilling SOC analysts in AI technologies and decision-making methodologies to leverage the new framework effectively. Engaging stakeholders and fostering buy-in throughout the implementation process is essential for driving successful adoption.

7.2.7.8.4 Risk Mitigation Strategies

Transitioning to an AI-driven decision-making framework inevitably introduces new risks that demand careful consideration and proactive management. One critical aspect involves addressing algorithmic biases, where thorough testing and validation procedures can help identify and mitigate potential biases in the AI algorithms. This could include diverse dataset sampling, regular audits, and continuous monitoring to ensure fair and equitable outcomes. Moreover, fostering transparency by documenting the decision-making processes of AI models and making them understandable to stakeholders can enhance trust and accountability, thereby mitigating risks associated with opacity and lack of explainability.

Another pivotal aspect of risk mitigation involves safeguarding data privacy and security. Implementing robust data governance frameworks, including encryption, access controls, and anonymization techniques, can help protect sensitive information from unauthorized access or breaches. Furthermore, adhering to regulatory compliance standards such as General Data Protection Regulation (GDPR) or California Consumer Privacy Act (CCPA) ensures that data handling practices are aligned with legal requirements, reducing the likelihood of regulatory penalties and reputational damage. Additionally, establishing mechanisms for ongoing monitoring and evaluation of AI systems can facilitate the timely identification and mitigation of emerging risks, enabling adaptive responses to evolving threats and challenges in the AI-driven decision-making landscape.

7.2.7.8.5 Financial Consideration

As part of our AI strategy, we recognize the need for financial investment to support AI initiatives. To facilitate the computational infrastructure required for the proposed framework, we propose releasing a new budget allocation of 600K SG. This investment will ensure that we have the necessary resources to support AI-driven decision-making effectively.

7.2.8 AI-Human Collaborative Decision Making

To address these flaws, an AI-human collaborative decision-making framework is proposed for SOCs. Collaborative intelligence between humans and AI enhances decision-making efficiency and facilitates smoother execution of tasks. While AI excels in quantitative measurements, human involvement is essential in interpreting ethical standards and ensuring decisions align with broader moral imperatives Section 7.2.7; Kim (2023).

7.2.9 AI-Driven Decision Making and Problem Solving Processes

AI-driven alert generation and triage processes in SOCs utilize machine learning algorithms to analyze incoming alerts and provide actionable insights to analysts. These processes combine automated analysis with human oversight to prioritize alerts, categorize them, and recommend response actions. Additionally, reinforcement learning algorithms enable continuous improvement by analyzing historical data and identifying patterns and trends for optimization Davianto (2022); MITRE (n.d.).

Continuous monitoring of the organization's network, systems, and applications by various security tools and sensors initiates the alert generation process. These tools detect anomalies or suspicious activities, generating numerous security alerts based on predefined rules. To combat challenges like alert fatigue and cognitive biases, AI-driven algorithms facilitate initial filtering and prioritization, combining automated analysis with human oversight by SOC analysts.

AI plays a pivotal role in scrutinizing individual alerts and cross-referencing them with other alerts within the system, utilizing machine learning techniques to discern patterns, anomalies, and relationships among alerts. By considering various factors such as signature and temporal-based characteristics, recent security events, threat intelligence feeds, and historical data, AI provides valuable insights into the potential severity and relevance of alerts. Moreover, it offers explanations for alerts, assisting SOC analysts in their investigative endeavors.

Transitioning from alert generation to alert triage, sophisticated machine learning algorithms and AI-powered decision support tools autonomously analyze incoming alerts. They utilize contextual information from the alert generation step, threat intelligence feeds, historical data, and real-time insights to assess severity, relevance, and potential impact on the organization's security posture. Leveraging vast repositories of historical data and threat intelligence, AI contextualizes incoming alerts, categorizing them into different tiers or risk levels.

Furthermore, AI-powered decision support tools provide Tier one and two analysts with actionable recommendations, including steps for further investigation, response actions to mitigate threats, and additional data sources for deeper analysis. The triage process incorporates mechanisms for continuous learning and improvement, enabling AI algorithms to adapt and refine their decision-making capabilities over time. By integrating frameworks such as MITRE (n.d.), AI enhances its ability to detect known attack patterns and techniques, correlating observed alert characteristics with known attack patterns for more accurate threat detection and attribution. Through this integration, AI streamlines the entire alert management process, improving efficiency, reducing response times, and enhancing the overall security posture of the organization.

In the context of cybersecurity, the utilization of AI-driven automation is not merely a futuristic concept but a pressing necessity to address the evolving landscape of cyber threats and the increasing complexity of security operations. Building upon the foundations laid out in the previous sections, let's delve deeper into how AI will reshape and optimize the decision-making processes within Security Operations Centers (SOCs).

7.2.10 Continuous Improvement

AI-driven approaches, bolstered by reinforcement learning, foster a culture of perpetual learning and innovation in SOCs. By analyzing historical incident records and response efficacy metrics, AI algorithms identify patterns and trends, offering opportunities for improvement. Clear metrics and key performance indicators (KPIs), such as reduction in noise within decision-making and time from threat detection to mitigation, enable organizations to assess the effectiveness of AI-driven decision-making frameworks Kim (2023); MITRE (n.d.).

7.2.11 Proposed AI-Human Collaboration Paradigm

In the envisioned paradigm of AI-human collaboration within Security Operations Centers (SOCs), human managers orchestrate a network of AI agents to handle various SOC functions. These AI agents, operating collectively as a swarm, manage alert generation, triage, incident response, and documentation tasks, communicating seamlessly with humans and each other through natural language Shambaugh (2018). Human oversight guides AI evolution and performance tuning based on established KPIs and metrics, ensuring continuous improvement De Smet (n.d.).

7.2.12 Alignment with Identified Flaws

The proposed data-driven decision-making approach for Security Operations Centers (SOCs) addresses various flaws inherent in traditional decision-making processes.

7.2.12.1 Alert Fatigue and Cognitive Biases

The approach mitigates alert fatigue and cognitive biases by leveraging AI-driven algorithms for alert generation and triage. Automation reduces the overwhelming volume of alerts, minimizing the risk of fatigue and oversight. AI provides objective assessments of alert severity, countering biases that may influence human judgment.

7.2.12.2 Accuracy and Efficiency

AI-driven analytics and machine learning enhance the accuracy and efficiency of threat detection. These models can process vast amounts of data, identify subtle patterns, and adapt to evolving threats, addressing the challenge of scarce expertise and dynamic cyber threats.

7.2.12.2.1 Rapid Response

Integration of AI-driven orchestration and automation tools streamlines workflows, facilitates rapid response actions, and enhances collaboration among response teams. By automating routine tasks, analysts can focus on high-value tasks like threat analysis and decision-making, mitigating the risk of errors or delays.

7.2.12.2.2 Documentation and Reporting

Adoption of AI-driven natural language processing (NLP) automates the generation of comprehensive incident reports and case summaries. AI-driven NLP algorithms improve the quality and consistency of documentation, enabling trend analysis and predictive modeling within incident reports.

7.2.12.2.3 Overall Enhancement

The data-driven decision-making approach leverages AI-driven technologies to enhance decision accuracy, responsiveness, and effectiveness, strengthening cybersecurity posture and resilience against evolving threats.

7.2.12.2.4 Noise Reduction in Decision Making

Continuous improvement through AI-driven methodologies and reinforcement learning mitigates bias and inconsistency in decision-making. Objective analysis of vast decision-making data facilitates actionable insights and process optimization, reducing the influence of noise and uncertainty.

7.2.12.2.5 Evaluation and Optimization

AI-driven analytics facilitate evaluation of individual performance and process effectiveness, pinpointing areas for optimization and resource allocation. Reinforcement learning approaches improve decision-making under uncertainty, optimizing performance in a dynamic cybersecurity landscape.

7.2.12.2.6 Integration with Identified Flaws

The proposed approach directly addresses challenges such as alert fatigue, cognitive biases, and the need for rapid response in SOCs. By leveraging AI-driven technologies, it enhances decision-making accuracy, responsiveness, and documentation quality.

7.2.12.2.7 Synergy with Organizational Objectives

Aligning with organizational goals, the approach promotes efficiency, effectiveness, and resilience in cybersecurity operations. It fosters collaboration, innovation, and continuous improvement, contributing to overall SOC performance.

7.2.13 AI-Human Decision Making

This collaborative framework recognizes AI as a supportive tool, leveraging its analytical prowess while valuing human judgment for nuanced decision-making. AI-driven algorithms act as decision support systems, providing data-driven insights and recommendations for human interpretation. Transparency, explainability, and accountability are prioritized, ensuring human oversight and alignment with ethical standards Kim (2023); De Smet (n.d.).

7.2.14 Continuous Improvement

The proposed approach emphasizes continuous improvement through reinforcement learning, enabling SOCs to adapt to evolving challenges and optimize performance. AI-driven analytics facilitate objective analysis of decision-making data, identifying patterns, trends, and areas for enhancement. By proactively recalibrating detection algorithms and response strategies, SOCs can preemptively address emerging threats, reducing noise and uncertainty Kahneman (2021).

7.2.15 Reinforcement Learning for Noise Reduction

Reinforcement learning algorithms enhance singular recursive and evaluative judgments within SOC operations. These algorithms guide decision-making under uncertainty by maximizing cumulative rewards, which encompass metrics measuring overall system performance. By iteratively refining decision-making protocols and adapting to evolving threats, SOCs can optimize performance in a dynamic cybersecurity landscape, mitigating the impact of noise and uncertainty Kahneman (2021).

8 References

Adams, J. (n.d.) The Five Whys.PDF *Google Docs*. [Online]. Available at: https://docs.google.com/document/u/3/d/1bqtDmkbdS7XISEvRa6L9AuJxTmillkr0ty5KCN8KIgE/edit?usp=drive_web&ouid=102557099045858632615&usp=embed_facebook (Accessed: 20 March 2024).

Christensen, C. (2013) *The innovator's solution: Creating and sustaining successful growth*. [Online]. Harvard Business Review Press. Available at: https://books.google.com.sg/books?id=r0xxJUzyFHYC.

Davianto, H. (2022) The Advantages of Artificial Intelligence in Operational Decision Making. *Hasanuddin Economics and Business Review*. [Online] 6 (1), 24. Available at: doi:10.26487/hebr.v6i1.5082 (Accessed: 10 March 2024).

De Smet, A. (n.d.) *How boards can make better decisions | McKinsey* [Online]. Available at: https://www.mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/boards-and-decision-making (Accessed: 20 March 2024).

Dimitrov, W. and Syarova, S. (2019) Analysis of the functionalities of a shared ICS security operations center. In: *2019 big data, knowledge and control systems engineering (BdKCSE)*. [Online]. 2019 pp. 1–6. Available at: doi:10.1109/BdKCSE48644.2019.9010607.

Garofalo, S. (n.d.) *Sara Garofalo: The psychology behind irrational decisions | TED Talk* [Online]. Available at: https://www.ted.com/talks/sara_garofalo_the_psychology_behind_irrational_decisions (Accessed: 12 March 2024).

Gilbert, C. and Bower, J.L. (2002) Disruptive Change: When Trying Harder Is Part of the Problem. *Harvard Business Review*. [Online] Available at: https://hbr.org/2002/05/disruptive-change-when-trying-harder-is-part-of-the-problem (Accessed: 13 May 2024).

Goel, V. (2010) Neural basis of thinking: Laboratory problems versus real-world problems. *WIREs Cognitive Science*. [Online] 1 (4), 613–621. Available at: doi:10.1002/wcs.71 (Accessed: 12 March 2024).

Hammer, M. (1990) Reengineering Work: Don't Automate, Obliterate. *Harvard Business Review*. [Online] Available at: https://hbr.org/1990/07/reengineering-work-dont-automate-obliterate (Accessed: 11 May 2024).

Inc, D. (2017) *Insights into building an industrial control system security operations center.technical report. 12 pages*

Ionescu, S.-A. and Diaconita, V. (2023) Transforming Financial Decision-Making: The Interplay of AI, Cloud Computing and Advanced Data Management Technologies. *INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL*. [Online] 18 (6). Available at: doi:10.15837/ijccc.2023.6.5735 (Accessed: 10 March 2024).

Jacq, O., Boudvin, X., Brosset, D., Kermarrec, Y., et al. (2018) Detecting and Hunting Cyberthreats in a Maritime Environment: Specification and Experimentation of a Maritime Cybersecurity Operations Centre. In: *CSNet 2018: 2nd Cyber Security In Networking Conference*. [Online]. October 2018 Paris, France. p. Available at: doi:10.1109/CSNET.2018.8602669.

Jesuthasan, R. (2019) The 8 Ways Companies Get Work Done, and How to Align Them. *Harvard Business Review*. [Online] Available at: https://hbr.org/2019/08/the-8-ways-companies-get-work-done-and-how-to-align-them (Accessed: 27 April 2024).

Jesuthasan, R. and Boudreau, J.W. (2018) *Reinventing Jobs: A 4-Step Approach for Applying Automation to Work*. Boston, Massachusetts, Harvard Business Review Press.

Kahneman, D. (2021) *Noise*. [Online]. Available at: https://www.hachettebookgroup.com/titles/daniel-kahneman/noise/9780316451390/ (Accessed: 17 March 2024).

Kim, H. (2023) Suggestions for Ethical Decision-Making Model through Collaboration between Human and AI. *Robotics & AI Ethics*. [Online] 12–22. Available at: http://scholar.kyobobook.co.kr/article/detail/4010060829725.

Kirsner, S. (2021) Don't Let Financial Metrics Prematurely Stifle Innovation. *Harvard Business Review*. [Online] Available at: https://hbr.org/2021/03/dont-let-financial-metrics-prematurely-stifle-innovation (Accessed: 27 April 2024).

Lean Enterprise Institute (2018) *Clarifying the '5 Whys' Problem-Solving Method* [Online]. Available at: https://www.youtube.com/watch?v=SrlYkx41wEE (Accessed: 12 March 2024).

LeCun, Y. (2022) A Path Towards Autonomous Machine Intelligence *OpenReview*. [Online]. Available at: https://openreview.net/forum?id=BZ5a1r-kVsf (Accessed: 8 May 2024).

MITRE, C. (n.d.) *MITRE ATT&CK®* [Online]. Available at: https://attack.mitre.org/ (Accessed: 22 March 2024).

Onwubiko, C. and Ouazzane, K. (n.d.) *Challenges towards Building an effective Cyber Security Operations Centre – c-mric.com* [Online]. Available at: https://c-mric.com/100124 (Accessed: 16 April 2024).

Podolny, J.M. and Hansen, M.T. (2020) How Apple Is Organized for Innovation. *Harvard Business Review*. [Online] Available at: https://hbr.org/2020/11/how-apple-is-organized-for-innovation (Accessed: 19 March 2024).

Sankaran, G., Palomino, M.A., Knahl, M. and Siestrup, G. (2022) A modeling approach for measuring the performance of a human-AI collaborative process. *Applied Sciences*. [Online] Available at: https://api.semanticscholar.org/CorpusID:253664653.

Shambaugh, R. (2018) Managing Someone Who's Too Collaborative. *Harvard Business Review*. [Online] Available at: https://hbr.org/podcast/2018/08/managing-someone-whos-too-collaborative (Accessed: 20 March 2024).

Wujec, T. (n.d.) *Tom Wujec: Got a wicked problem? First, tell me how you make toast | TED Talk* [Online]. Available at: https://www.ted.com/talks/tom_wujec_got_a_wicked_problem_first_tell_me_how_you_make_toast (Accessed: 12 March 2024).

9 Document Statistics

Total word count: ***8662***