

# **A Framework for Improving Intrusion Detection Systems by Combining Artificial Intelligence and Situational Awareness**

DANIEL LINDEMANN & YAGUEL VAN DER MEIJ

**SUPERVISOR**  
Paolo Spagnoletti

University of Agder, 2022  
Faculty of Social Sciences  
Department of Information Systems



# PREFACE

This thesis is written as a final assignment for the master's degree in Cybersecurity - Security Management at the department of Social Sciences at the University of Agder and was conducted from January 2022 to June 2022.

The thesis is conducted in order to evaluate and analyze the state-of-the-art in automated security solutions, and how this can be improved. Additionally, the thesis aims to provide a framework that can be used in future research on this topic, as well as incorporated by businesses that aim to improve their cybersecurity operations.

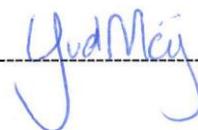
The motivation for selecting this topic is founded in emerging trends in the cyber-threat-landscape that correlates with trends in digitalization.

Firstly, we would like to thank our supervisor, Paolo Spagnoletti, for the valuable feedback and guidance he has provided for us during this thesis project. He has been easily reachable and always encouraging.

Lastly, we would like to thank the domain expert who has agreed to participate in interviews to help provide knowledge and insights into the topics of this thesis, as well as to help evaluate and provide feedback on our framework.

Kristiansand  
June 2<sup>nd</sup>, 2022  
Daniel Lindemann

Yaguel van der Meij



---



## ABSTRACT

The vast majority of companies do not have the requisite tools and analysis to make use of the data obtained from security incidents in order to protect themselves from attacks and lower their risk. Intrusion Detection Systems (IDS) are deployed by numerous businesses to lessen the impact of network attacks. This is mostly attributable to the fact that these systems are able to provide a situational picture of network traffic regardless of the method or technology that is used to generate alerts. In this paper, a framework is proposed for improving the performance of contemporary IDSs by incorporating Artificial Intelligence (AI) into multiple layers, presenting the appropriate abstraction and accumulation of information, and generating valuable logs and metrics for security analysts to use in order to make the most informed decisions possible. This is further enabled by including Situational Awareness (SA) at the fundamental levels of the framework.

**Keywords:** Intrusion Detection System, Machine Learning, Deep Learning, Shallow Learning, Security Operation Center, Situational Awareness

## Table of contents

1	INTRODUCTION.....	1
1.1	Intrusion Detection Systems.....	1
1.2	Situational Awareness.....	2
1.3	Security Operations Center.....	2
1.4	Significance.....	2
1.5	Research Question.....	3
1.6	Structure of the Report.....	3
1.6.1	Chapter 2: Background and Related Work.....	3
1.6.2	Chapter 3: Research Approach.....	4
1.6.3	Chapter 4: Artifact.....	4
1.6.4	Chapter 5: Discussion.....	4
1.6.5	Chapter 6: Conclusion.....	4
2	BACKGROUND AND RELATED WORK.....	5
2.1	Literature Findings.....	5
2.1.1	Machine Learning.....	6
2.1.2	Intrusion Detection System.....	8
2.1.3	Security Operation Center.....	9
2.1.4	Situational Awareness.....	10
2.2	Conclusion from the Theoretical Background.....	11
3	RESEARCH APPROACH.....	13
3.1	Design Science Research.....	13
3.2	DSR in our Project.....	16
3.2.1	Activity 1: Problem Identification and Motivation.....	16
3.2.2	Activity 2: Define Objectives for a Solution.....	16
3.2.3	Activity 3: Design and Development.....	17
3.2.4	Activity 4: Demonstration.....	17
3.2.5	Activity 5: Evaluation.....	17
3.2.6	Activity 6: Communication.....	18
3.3	Literature Review Methodology.....	18
3.3.1	Systematic Literature Review.....	18
3.3.2	The Purpose.....	19
3.3.3	Draft Protocol.....	20
3.3.4	Practical Screen.....	20
3.3.5	Search for Literature.....	20
3.3.6	Appraise Quality.....	21
3.3.7	Extract Data.....	21
3.3.8	Synthesis of Studies and Writing the Review.....	23
3.4	Interview with Domain Expert.....	25

4	ARTIFACT .....	26
4.1	Reference Framework .....	26
4.2	Our Proposed Improvements .....	28
4.2.1	Artificial Intelligence .....	28
4.2.2	Appliance Connection to Network .....	30
4.2.3	Event Log Handling .....	30
4.2.4	Situational Awareness .....	30
4.3	Our framework .....	31
4.3.1	The detailed phases .....	32
5	DISCUSSION .....	36
5.1	An AI-based framework for enhanced IDS in SOC .....	36
5.1.1	Machine Learning .....	36
5.1.2	Situational Awareness .....	37
5.1.3	Security Operation Center .....	37
5.1.4	Intrusion Detection System .....	38
5.2	Challenges & Limitations .....	38
5.3	Future research .....	38
6	CONCLUSION .....	40
	REFERENCES .....	41
	APPENDIX .....	47

## List of figures

Figure 1	DSR process applied to our research problem. Adapted from Peffers et al., (2007). .....	15
Figure 2	An automated framework for triage, containment, and escalation. Source: (Danquah, 2020).....	26
Figure 3	Improved framework for automated security using artificial intelligence and situational awareness. ....	32

## List of tables

Table 1	Literature overview .....	22
Table 2	Categorical overview of literature .....	24

# 1 INTRODUCTION

Organizations increasingly utilize the many benefits that come with the significant advances in digital technologies, such as increased efficiency, speed of communication, and accessibility (de Reuver et al., 2017). These advances assist organizations to gain a better position in the market and relieve laborious tasks so that focus can be placed elsewhere (Şerban, 2017).

With these advances in digitalization and digital technology use comes a growing risk of cyber-attacks. If a cyber-attack is successful, it can compromise the integrity, reliability, and confidentiality of data and services. Such attacks can often come in great volumes and can be very unpredictable (Y. Li & Liu, 2021). The increasing number of cyber-attacks accentuate the importance of the implementation of proper cyber security. Cyber security can be described as the procedure of protecting an organization's digital-related assets, which includes all the data that is managed within an organization's systems and networks connected to the Internet (Kim & Solomon, 2019).

## 1.1 Intrusion Detection Systems

Due to the number of attacks, manually observing network traffic and isolating malicious events is unfeasible, and automation needs to be introduced. A commonly used tool is Intrusion Detection Systems (IDS). An IDS is a software or hardware system that automates the process of monitoring the events occurring in a computer system or network and analyzing them for signs of security problems (Bace & Mell, 2001). The use of IDS was popularized in the early 2000s, but the technology is frequent practice today in most organizations in one form or another, often in conjunction with other systems to assist in the process (Pirc, 2017).

While IDSs are critical in detecting the vast number of threats, there are some issues. To be able to detect new threats, an anomaly base IDS also produces a large number of false positives that are time consuming to analyze and process. The alternative to this is a misuse base IDS, which only detects known threats (Almseidin et al., 2017). In addition to this, attacks are getting more sophisticated with supply-chain attacks becoming more common and harder to detect. This furthers the need for improvement in detection of threat events (Enisa, 2021).

## 1.2 Situational Awareness

Protecting information resources from sophisticated and persistent cyber-attacks is a crucial challenge. Organizational environments change rapidly and the socio-technical systems that include personal, software, hardware, and community aspects produce elevated levels of information with a lot of diversity. To be able to understand these socio-technical environments and anticipate what might happen while trying to predict potential threats requires a complete picture of all the systems, networks, processes, and users in order to develop Situational Awareness (SA) (A. Ahmad et al., 2021).

SA is a combination of three stages creating awareness on an operational, tactical, and strategic stage making sure that all aspects are taken into consideration. While operational awareness focuses on exposing the impact to its operations by correlating the context obtained from a tactical perspective. The tactical SA stands for the understanding of events and situations. The strategic stage focuses on being able to expose the malicious objective of the potential threat actors and recognize trends in their activities (MITRE, 2022).

## 1.3 Security Operations Center

A common way to address SA is by analyzing network traffic, firewall/IPS, and threat intelligence with the implementation of a Security Operation Center (SOC) (Ponemon Institute, 2020). A SOC is a centralized organizational unit that employs people, processes, and technology to monitor, collect, and analyze security events throughout an organization's IT infrastructure and security controls (McAfee, n.d.). SOCs are normally led by security specialists who manually analyze the alerts to conclude whether they truly are malicious. The challenge with analyzing the alerts a SOC produces is often the sheer number of false positives, making it hard to identify threats (Brewer, 2019; Gupta et al., 2019).

## 1.4 Significance

With the rise of digitalization and the expansion of the cyber-threat landscape, it is evident that this is a problem that numerous organizations are coping with and improvements are sought after by many (Ahmim et al., 2019; Bringhamti et al., 2019; Kasongo & Sun, 2019; Y. Li & Liu, 2021; Mishra et al., 2019; Naseer et al., 2018; Shone et al., 2018; Vast et al., 2021; Zeadally et al., 2020). Until now, academics from all around the world have proposed a variety of ways to prevent cyber-attacks or reduce the false alarm rate and the majority are geared toward

finding a solution to difficulties in highly specialized sectors, whether technological or administrative.

There is a lack of a framework which proposes combining different layers of Machine Learning (ML) models to create a holistic solution that collects data from various algorithms to provide SA for security analysts to make the best decisions. That is why we propose a framework that improves the performance of current IDSs by incorporating Artificial Intelligence (AI) into several layers, presenting the appropriate abstraction and accumulation of information to compensate for human perception and cognition limitations, and that can be used to generate valuable logs and metrics for security analysts to make the best decisions.

## **1.5 Research Question**

To address the above-mentioned research gap, we focus on the following research question:

- How can AI and SA be successfully integrated to improve the performance of IDSs?

To answer this question, we modify and extend an existing framework proposed by Danquah (2020). This extended framework combines several methods for improving cyber security automation as well as providing a situational picture of network traffic. The framework provides a holistic view of IDS operations that relate to a SOC, detailing the flow of activities and separating them in groups of automatic and manual processes.

## **1.6 Structure of the Report**

The remainder of the chapters of this thesis are described in this section, as well as an overview of their contents.

### ***1.6.1 Chapter 2: Background and Related Work***

This chapter covers previous research and literature that is relevant to our research project. In this chapter we describe concepts and theories regarding ML, IDSs, SOCs, and SA.

### ***1.6.2 Chapter 3: Research Approach***

This chapter describes the methodologies used when conducting our research. Firstly, we present the research approach and methodology, with a detailed walkthrough of the method. Then we present the method for conducting our literature review, as well as the steps of our overall research process.

### ***1.6.3 Chapter 4: Artifact***

This chapter describes the development of our artifact. We start by explaining the reference framework we extended and modified, as well as detailing how we aimed to improve it and what issues we found. We then present our version of the framework and explain it thoroughly.

### ***1.6.4 Chapter 5: Discussion***

In this chapter we discuss our findings, our approach and methodology, as well as limitations of our thesis and future works.

### ***1.6.5 Chapter 6: Conclusion***

In this last chapter, we summarize the research goal, the research method, the findings, and contribution.

## 2 BACKGROUND AND RELATED WORK

The following section provides an overview of previous research that is relevant to this thesis. A systematic literature review has been conducted to ascertain the state of the art in the area of cyber security intrusion detection. This helped to facilitate the development of our solution while gaining knowledge on existing research and find areas where research is needed. The details on our approach and strategies are covered in more depth later.

This thesis aims to expand on the state-of-the-art methodologies used in intrusion detection by gaining a better understanding of the technical aspects, as well as the organizational strategies that are available at the time of writing this thesis.

### 2.1 Literature Findings

After narrowing down the papers that were relevant to our research, we tried to synthesize the literature we read by examining the key concepts. The most common concept in the literature we found is the use of Deep Learning (DL), a complex area of ML, to aid in anomaly detection in IDS. This topic was covered extensively in nineteen of the papers we focused on. There were some varying focus areas within the topic of DL in these papers that we will cover in the following section specific to DL.

Another topic that is prevalent is Shallow Learning (SL), the opposite of DL. Eight of the articles propose the use of SL to increase the accuracy in incident detection of unknown events. This topic was often coupled with DL, but some papers exclusively covered SL.

Some papers focused on the general aspects of IDS and how they may be improved. This also included papers covering Network-IDS (NIDS) and Anomaly-based IDS (AIDS). Some of these papers mention the use of SOC, ML, DL, etc. but do not focus on them, which is why we have separated them into a general category of IDS.

Furthermore, there are some articles that have investigated processes and practices taking place within SOCs and issues that are arising there. Some focus on triage and the way forward with improving automated triage in SOC, and some focus purely on automation within SOC. One paper also looks at the human aspect

of SOC and how the human aspect is important when designing automated solutions.

There were also a couple of papers that focused on User and Entity Behavior Analytics (UEBA). This is a subset of ML that focuses on mapping the “normal” behavior of users and entities on a network to isolate events where users and entities act abnormally.

Despite the focus on advanced techniques and technical components to support SOC activities, a socio-technical view of AI-based solutions can improve the effectiveness of security operations. As a result, we have looked for articles that focused on implementing SA using intrusion detection.

### ***2.1.1 Machine Learning***

Over half of the papers, we collected discuss and research the use of ML in IDS. We want to separate the common methods within ML as the lines can easily be blurred. The most common methods of ML are SL, the simplest form of ML, and DL, a more complex and comprehensive form of ML.

#### *Shallow Learning*

The biggest benefit of utilizing ML is the ability to automatically detect unknown attacks, which is why several papers mention that using SL models can be especially useful since they are easier to design and construct compared to DL models (Gümüþbaþ et al., 2021; Liu & Lang, 2019; Xin et al., 2018). Other advantages to SL models that are mentioned by Liu & Lang are the shorter running time for both training and testing, the lower number of parameters making it faster to optimize and train, and they require less volume of data to be trained.

Xin et al. (2018) describe that SL learning performs better as compared to DL models when the data volumes are small. Another factor that is mentioned is that a computer running SL algorithms needs less high-performance hardware compared to what is required by DL algorithms. They also mention that the results generated from SL models are easily interpreted since they follow exact rules which explain why choices are made.

Shaukat et al. (2020) provide an extensive review of several ML and DL models in which they describe how the chosen models can have been or can be used to detect and classify cyberattacks in various kinds of tools such as IDSs. They also state that DL models need more data to perform well and that SL models are beneficial when the amount of data is insufficient.

### *Deep Learning*

The measures for improving IDS that we find as most prevalent is DL and ML. While they are similar, there are enough differences that we want to separate them.

When it comes to improving the detection of unknown and unanticipated attacks, the most suggested asset is DL. Several reports suggest the use of various Deep Neural Network (DNN) models to develop a flexible and effective IDS to detect and classify unforeseen and unpredictable cyberattacks (Z. Ahmad et al., 2021; Aldweesh et al., 2020; Berman et al., 2019; Dixit & Silakari, 2021; Kasongo & Sun, 2019; Liu & Lang, 2019; Shone et al., 2018; Vinayakumar et al., 2019; Wang, 2018).

Mahdavifar and Ghorbani (2019) state that it is important that DL should not be used in every domain. Instead, they say that DL should be utilized in area of high complexity such as non-linear hypotheses with many features and high-order polynomial terms and in domains with large-scale data, hence an IDS, helping to reduce the heavy processing of the input data. Another article which uses a DL approach also shows promising results. In their paper they claim to have achieved higher precision with an average accuracy of around 0.918 (Moraboena et al., 2020). The survey of Xin et al. (2018) explains a variety of SL and DL models with IDS as a focus, and they add that the hybrid method, which combines both SL and DL, has received less research but is quite promising. Another research uses graphical analysis and classification to provide a new strategy to implementing DL within a SOC. By recognizing and turning important characteristics and relationships into new features (Gupta et al., 2019).

Apruzzese et al. (2018) conducted research to investigate which algorithm works better in various cyber-related scenarios. They compare the performance of Random Forest (SL) and Feedforward Fully Connected Deep Neural Network (DL) in terms of improving Domain generation Detection and NIDS. Their findings show that the SL model outperforms the DL model in some cases, and they conclude that more research is needed.

Rahul et al. (2018) compare deep neural networks with a variety number of layers, as well as SL models, in their paper. For benchmarking, they largely employed the KDDCup-'99 dataset, and their article indicates encouraging results for deep neural networks in cybersecurity.

Rahul et al. (2018) compare deep neural networks with a variety number of layers, as well as SL models, in their paper. For benchmarking, they largely employed the KDDCup-'99 dataset, and their article indicates encouraging results for deep neural networks in cybersecurity.

Even so, several articles also point out that future research should focus on minimizing training times and that many of the public training datasets are outdated, contain a small number of samples and redundant records (Ferrag et al.,

2020; Karatas et al., 2019; Liu & Lang, 2019; Mahdavifar & Ghorbani, 2019; Wang, 2018; Xin et al., 2018).

### *User and Entity Behavior Analytics*

One paper proposes to use UEBA to improve security of Federated Identity Management (FIM) solutions. The proposed solution allows the creation of fingerprints characterizing each user's behavior from available information. This enables anomaly detection based on the fingerprints (Martín et al., 2021).

Another paper aimed to highlight weaknesses and strengths of different UEBA solutions and their effectiveness for detecting attacks in real-time interaction. They compare fifteen of the top UEBA technologies based on use cases and technologies and highlight common scenarios of use (Salitin & Zolait, 2018).

In addition, UEBA is mentioned in a report about Security Orchestration and Automated Response (SOAR) systems by Kinyua & Awuah (2021).

### **2.1.2 Intrusion Detection System**

A few articles research IDS in general, and map the current state of the art, find weaknesses and list different types of IDS.

Khraisat et al. (2019) performed a survey of IDS techniques, datasets, and challenges where they present detailed information about these different topics as well as common attack strategies utilized against IDS and how future research can focus on the mitigation of such attacks.

Another paper provides an overview of existing articles that focus on incident prediction instead of detection (Sun et al., 2019). They describe it as a system being able to proactively act against unknown threats while providing knowledge on how to improve security instead of just detecting and mitigating threats. Their research is divided by the common steps of a data-driven research methodology and in their research, they state that improving overall security by giving social or financial incentives may be more efficient than developing new technological solutions. They also mention that using representation learning to identify unknown factors could make significant advances. Another point they bring up is that careful selection of natural language processing tools and customization for specific domains is important to achieve higher performance.

### ***2.1.3 Security Operation Center***

Several papers suggest utilizing a SOC with improvements to enhance the ability to manage vulnerabilities, risks, and security incidents by monitoring, responding, preventing, and reporting security related events (A. Ahmad et al., 2021; Danquah, 2020; Gupta et al., 2019; Lin T et al., 2018; Zhong et al., 2019). According to the articles it is critical for every organization to have a SOC in place to make sure any potential threat is flagged as early as possible.

#### *Issues*

Utilizing a SOC might strengthen security, though considering the sheer amount of data that is produced daily, it is obvious that this can also make it challenging to keep up with the workload. SOCs are normally led by a security specialist who manually analyzes the alerts to conclude whether or not they truly are malicious and as many of the papers describe, the challenge with analyzing the alerts a SOC produces is often the sheer number of false positives and false-negatives, making it hard to identify threats (Chamkar et al., 2021; Gupta et al., 2019; Lin T et al., 2018; Mohsienuddin Mohammad & Lakshmisri, 2018; Vielberth et al., 2020).

The paper of Vielberth et al. (2020) mentions several challenges that come with the use of a SOC. They state that there is a lack of skilled staff due to insufficiency in job-related security training and that the knowledge of analysts is rarely shared among one another. Another issue they describe is the fact that the current ML algorithms are well trained for known attacks, but it is still challenging to train algorithms on unknown factors.

#### *Triage*

Danquah (2020) proposes a framework that contains eight steps to efficiently perform triage of security threats, vulnerabilities, and incidents, effectively contain identified breaches and appropriately escalate for prompt and accurate solutions. The solution provided in this context is via the containment component once an accurate triage has been affected.

Another solution to minimize the amounts of noise in raw data has been developed by Zhong et al. (2019). They propose a graph-based trace mining approach which focuses on three insights to construct useful patterns for data triage. Namely, claiming that it is possible to do tracing of the analysts' discreetly and then mine the traces of data triage in an automated way to acquire the best components for complicated rules. Finally using these components to enable analysts to focus on even more complicated rules instead of starting from scratch. According to their result it shows that it is both feasible and more performant to conduct automated data triage by using the traces produced by analysts', and that it will reduce the number of false positives.

Lin et al. (2018) reviewed several existing retrieval methods regarding data triage to support junior analysts in their performance. Their paper focuses on solutions that bring forth the experts' knowledge from past data triage operations by looking into algorithms that match with knowledge retrieval systems. After discussing existing rule-based and context-based retrieval systems one of their suggestions is the use of a similarity measure to compare centroids of two graphs which could retrieve the analytical reasoning process, underlying logic, and reasoning strategies of an analyst in a time efficient manner. They further argue that the use of ML, more specifically, recurrent neural networks (RNN), might play an essential part in speeding up the process of data triage.

### *Automation*

Many if not most of the articles that focus on improving a SOC or SOAR argue that the use of AI/ML can make a substantial difference (Danquah, 2020; Kinyua & Awuah, 2021; Lin T et al., 2018; Zhong et al., 2019).

One of the papers suggests the use of some of the latest algorithms of deep reinforced learning (DRL) since they are well known to solve complex problems with high performance (Kinyua & Awuah, 2021). Further recommending a DRL system that utilizes multiple interacting agents as a large-scale solution.

The framework of Danquah (2020) mentioned earlier utilizes an algorithm that focuses on gathering data from all the relevant layers of a network and analyzes that data to see whether a potential threat can be mitigated based on a built-in remediation capability list. A disadvantage here is that the algorithm needs a list to be able to perform and this list will never be able to contain all the needed capabilities.

The method Zhong et al. (2019) provides also uses automation to reduce the number of false-positives and false-negatives. They used automation to mine the traces found while performing data triage to try and automate the process.

#### **2.1.4 *Situational Awareness***

Masduki et al. (2017) propose an IDS Metrics Framework for cyber situational awareness that, according to them, includes the most up-to-date technologies and approaches for data evaluation. The data is analyzed and compared to one or more reference points to reach a conclusion that establishes SA by issuing early warnings and directing follow-up actions, as well as effective preventive measures or mitigation strategies, which could be beneficial in a cyber situational awareness system's decision-making process.

According to the literature study of A. Ahmad et al. (2021), SA is a fundamental attribute of organizational incident response, and they describe the human

decision-making aspect necessary to traverse the intricacies of incident response in a dense socio-technical context. They describe SA from a technology standpoint as a process of gathering relevant and valuable data, fusing essential aspects of the data together, and obtaining insights from the merged data. By designing an information processing network, they demonstrate how organizations can control information flow to develop SA by identifying sources of information and expertise to build communication pathways and linkages between stakeholders. As a result, they built an incident response process model that focuses more on managerial practices and highlights how organizations can practice SA of the cyber-threat landscape and the larger business context, arguing that this model demonstrates how structural constraints in mature and sophisticated response capabilities can be addressed.

Fauri et al. (2019) propose a role-based system that monitors BACnet building automation networks and uses SA and intrusion detection to discover devices, classify them according to functional tasks, and detect deviations from the assigned roles. They claim that implementing SA and intrusion detection improves the understandability of alerts and the system's adaptability by observing, parsing, and interpreting network messages and extracting useful information about devices by building a network map to provide operators with details about their system, making it easier to detect attacks.

## **2.2 Conclusion from the Theoretical Background**

To limit the risk of cyber-attacks, organizations increasingly attempt to establish appropriate technical and operational countermeasures. However, the vast majority lack the necessary tools and analysis to appropriately manage these attacks and reduce risk by utilizing the data obtained from security events. Analyzing network traffic, firewall/IPS, and threat intelligence with the deployment of a SOC is a common technique to address these issues. Having a centralized solution that monitors, collects, and analyzes security events across an organization's IT infrastructure and security controls using people, procedures, and technology is a practical solution. As a result, IDSs are being utilized as one of the security components in a centralized solution to limit the impact of network attacks since they can provide a situational picture of network traffic regardless of the methodology or technology used to generate alarms. These IDSs frequently produce a large number of false alarm rates, motivating a large amount of research into how or whether ML can be applied. Both SL and DL have been proven to be effective techniques for extracting meaningful information from network traffic and forecasting normal and abnormal activities based on learnt patterns and

combining the two models would make it highly advantageous to gain from the capabilities that each model has to offer.

Organizational contexts change quickly, and socio-technical systems that combine personal, software, hardware, and community aspects generate a large amount of data with a wide range of content. To establish SA and anticipate what might happen in these socio-technical contexts while attempting to foresee potential risks, a full picture of all the systems, networks, processes, and users is required. To improve the decision-making process, the data created by the aforementioned ML models should be used to establish SA by delivering early warnings and guiding follow-up actions, as well as effective preventive measures or mitigation methods.

The majority of the research on the aforementioned topics were focused on finding a solution to tackle problems in very specialized fields, whether technological or managerial. We could not find anyone who suggested combining different layers of ML models to create a holistic solution that collects data from different algorithms and compensates for human perception and cognition limits to offer SA for security analysts to make the best decisions.

## 3 RESEARCH APPROACH

We decided to conduct a design science research (DSR) strategy for our project. This is a research strategy that is increasingly accepted and adopted by Information Systems Research (Deng & Ji, 2018).

### 3.1 Design Science Research

We use design science research in this thesis. In DSR research, artifact is designed and developed as a product of the project in this methodology. This method was chosen because it produces a product that is both beneficial and potentially valuable to the industry. This thesis' purpose was to develop a model for better incident anomaly detection in IDS and SOC.

For the project, we employed a series of activities based on a design science research methodology as described by Peffers et al. (2007). These activities are described below.

#### **Activity 1: Problem identification and motivation**

Define a specific research problem and justify the value of a solution. Because the problem definition will be used to design an artifact, it can be beneficial to break up the problem conceptually so that the complexity can be captured appropriately. This helps motivate the researcher and the audience.

#### **Activity 2: Define objectives for a solution**

From the problem definition and knowledge of what is achievable and practical, infer the solution's goals. The goals can be quantitative, such as the terms in which a desirable solution would be better than present ones, or qualitative, such as a description of how a new artifact is intended to assist answers to problems that have not been addressed previously. The objectives should be logically deduced from the problem description.

#### **Activity 3: Design and development**

Create the artifact. Such artifacts could be constructs, models, methods, or instantiations (all of which are widely defined) or new attributes of technical, social, and/or informational resources. A design research artifact can be defined as any created thing that incorporates a research contribution. This activity entails identifying the desired functionality and architecture of the artifact, as well as building the actual artifact. Knowledge of theory that can be applied to a solution

is one of the resources required for shifting from objectives to design and development.

#### **Activity 4: Demonstration**

Demonstrate the use of the artifact to solve one or more instances of the problem. This could include using it in experiments, simulations, case studies, proof, or other appropriate activities. The demonstration requires effective understanding of how to use the artifact to address the problem.

#### **Activity 5: Evaluation**

Examine and quantify how well the artifact contributes to a solution to the problem. This task involves comparing the goals of a solution to the actual results obtained from using the artifact in the demonstration. It requires familiarity with applicable measurements and analysis methodologies. Evaluation can take different shapes depending on the nature of the problem venue and the artifact. It could involve things like a comparison of the artifact's functionality to the solution objectives from activity two, objective quantitative performance indicators like budgets or items produced, satisfaction surveys, client feedback, or simulations, and so on. It could include quantifiable system performance measurements like response time or availability.

In theory, such an assessment may incorporate any suitable factual evidence or logical proof. At the end of this activity, the researchers can choose whether to iterate back to activity three to try to increase the artifact's effectiveness or to move on to communication and leave further improvement to future projects. The nature of the research site may determine whether such iteration is possible.

#### **Activity 6: Communication**

When applicable, communicate the problem and its importance, the artifact, its utility and originality, the rigor of its design, and its efficacy to researchers and other relevant audiences such as practicing professionals. Researchers may use the structure of this process to structure the paper in scholarly research publications, just as the nominal structure of an empirical research process (problem definition, literature review, hypothesis development, data collection, analysis, results, discussion, and conclusion) is a common structure for empirical research papers. Communication requires understanding of the disciplinary culture.

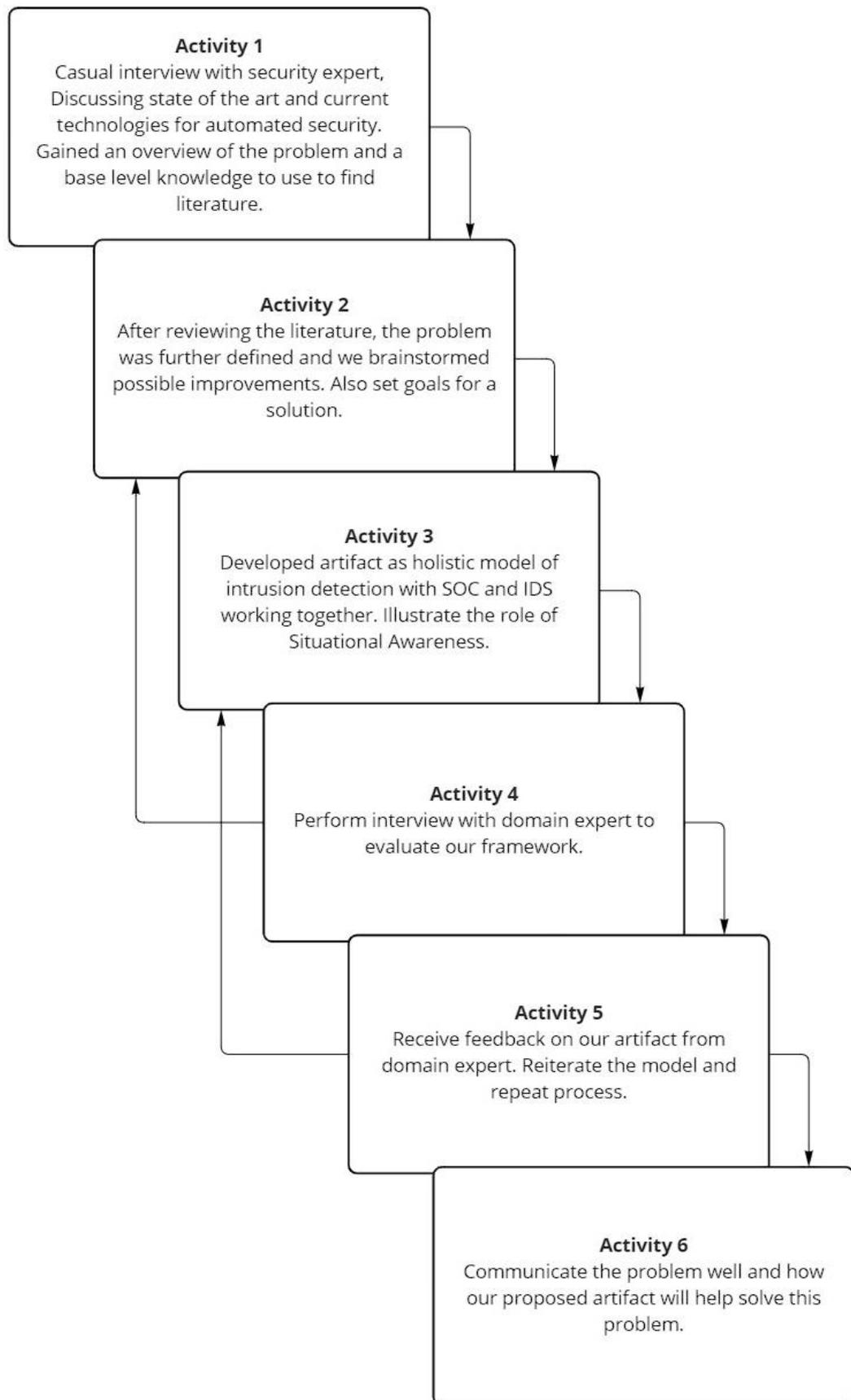


Figure 1 DSR process applied to our research problem. Adapted from Peffers et al., (2007).

## **3.2 DSR in our Project**

In Figure 1 shown above, we illustrate our process when following this model. In addition, we detail our process in each step below.

### ***3.2.1 Activity 1: Problem Identification and Motivation***

At the start of our project, we had an idea for what we wanted to research but did not have any specific goals or target area. We had the initial idea to look at automated security with a focus on cloud services. During an open-ended interview with an expert in the field of cybersecurity, we discussed the different aspects of cloud security, security automation and state of the art on these. We realized that while cloud security is an area which needs a more robust standard for security, the problem was not with cloud security, but incident detection and the problem that arises with false positive alerts.

In the latter part of the interview, we discussed intrusion detection and automation of response, and how false positives and false negatives impact the autonomy of a security system, as analysts need to monitor everything in parallel to be able to catch threats. Ideally, the system should be able to work autonomously most of the time, with little need for human interaction. This may be an unreachable goal for the foreseeable future, but improvement can be done, and this is the area we chose to focus on for this thesis.

The interview provided us with a lot of information and knowledge that we further used when gathering papers and articles for our literature review.

### ***3.2.2 Activity 2: Define Objectives for a Solution***

After having performed a review of previous research regarding IDS, SOC, AI in automated security, and SA, we analyzed and discussed the papers. We wanted to define a set of goals or “criteria” for what a successful solution to the problem should accomplish and how it should perform. We agreed that the solution should aim to reduce the number of false positives. In addition, we wanted to look at the human aspect with regards to intrusion detection, as humans are regarded as the weakest link in cybersecurity (Chamkar et al., 2021). To accomplish this, ML and SA emerged as two fundamental elements of our solution. ML, both shallow and deep, would help monitor different aspects of the network traffic and ideally provide a robust warning system with acceptable numbers of false positives. In addition, a focus on SA would provide the analysts with knowledge and tools to better collaborate with and understand the automated systems.

### ***3.2.3 Activity 3: Design and Development***

Once the overall goals of the solution have been identified, we went back to the previous research to see if anyone else had a similar idea. During this process, we came across a report by Danquah (2020). He had proposed a model using DL to reduce the number of false positives and increase autonomy in a SOC. His approach outlines technological processes that should be included in a system that attempts to effectively cover security threats, vulnerabilities, and events while also demonstrating that identified breaches are contained and appropriately escalated. The diagram shows how the various phases are implemented and how they interact. All of these are features we wanted to include in our framework. which helps to explain why we chose to use parts of it.

As a result, we decided to adopt this model as the basis for our own, and our goal was figuring out how we could show our ideas for improving automation and how SA could be implemented effectively.

To improve the model, we wanted to add SL as well as DL. This can be beneficial to handle simple ML tasks that DL would be excessive for, this would reduce the resources used by the system, thus ideally improving efficiency slightly (Mahdavifar & Ghorbani, 2019; Xin et al., 2018).

In addition to SL, we wanted to address SA (A. Ahmad et al., 2021). This is not a technical element, but an organizational one. SA in cybersecurity focuses on knowledge about security and threats, as well as how to address them. This aspect is important not only for the security analysts, but also the leaders of an organization, as this would help the leaders to understand how different situations can impact their organization and how many resources would be needed to help solve a critical situation.

### ***3.2.4 Activity 4: Demonstration***

In order to demonstrate our solution, we performed another interview with the same security expert as earlier. We then presented our model, explained the different elements and our ideas for how this could be used in practice. This led to a fruitful discussion about the state of things today, and how our different propositions can influence the efficacy of automated security solutions.

### ***3.2.5 Activity 5: Evaluation***

The evaluation of our model was achieved through the same interview that was conducted with the security expert. After providing a thorough demonstration and

discussion of the model, he then offered his comments. He made several observations on the efficacy of such a solution in general and offered some insightful feedback on particular aspects of our framework, both of which lead to a number of refactors.

### **3.2.6 Activity 6: Communication**

The significance of the problem that we have found is supported by both the review of the relevant literature and an interview with a domain expert who shared many of the same opinions as the general consensus that was found in the relevant literature.

## **3.3 Literature Review Methodology**

Following is a description of the methods used when conducting our literature review. The preferred literature review method for this report was systematic literature review (SLR).

### **3.3.1 Systematic Literature Review**

The systematic literature review method we used in this thesis was conducted as closely as possible according to the eight-step guide described by Okoli & Schabram (2012). They present a comprehensive guide that extends the fundamental methodology of several disciplines to meet the methodologically complex field of Information System research. Their guide aids researchers in identifying, evaluating, and synthesizing the existing body of previous significant work to ensure that a literature review is performed in a rigorous and beneficial way.

We chose to follow their eight-step guide as strictly as feasible in accordance with the scope of our thesis since their study focused more on conducting a standalone literature review.

Following is a description of their guide in accordance with our scope:

- 1. Identify the purpose:**

Clearly describing the review's aim and intended goals, to be explicit to the readers.

- 2. Draft protocol and train the team:**

Reviewers must be entirely clear and in agreement on the approach they will use to ensure consistency in how they execute the review.

**3. Apply practical screen:**

This step requires that the reviewers be explicit about what studies they considered for review, and which were dismissed without further investigation. Reviewers must justify how the resulting review can still be thorough given the practical exclusion criteria for excluded research by stating their practical reasons for not examining them.

**4. Search for literature:**

The reviewers must be detailed in describing the specifics of the literature search, as well as explain and justify how they ensured that the search was comprehensive.

**5. Appraise quality:**

The reviewers must state exactly whatever criteria they employ to determine which papers they will reject due to insufficient quality. Researchers must grade the quality of all papers included in the study based on the research criteria they employ. (In Okoli and Schabram's eight-step guide, this step comes after step 6, but because it did not make sense for us to extract data from papers that might have been removed after further review, we switched the order of steps five and six).

**6. Extract data:**

After reviewers have selected all the studies that should be included in the review, they need to systematically extract the applicable information from each study.

**7. Synthesize studies:**

This step involves combining the facts extracted from the studies by using appropriate techniques, whether quantitative, qualitative, or both.

**8. Write the review:**

A systematic literature review's process must be reported in sufficient detail so that other researchers can independently reproduce the review's result, in addition to the standard principles to be followed in writing research papers.

### ***3.3.2 The Purpose***

The main goal for this literature review was to gain a better understanding of the technical aspects of existing cloud-based IDS systems and methods, as well as the organizational strategies that were accessible at the time of writing this thesis. This realization has provided the knowledge needed to develop a framework that is better suited for organizations looking to boost security while reducing labor.

### 3.3.3 *Draft Protocol*

During the development and execution of the protocol to conduct the systematic literature review, the previously described eight steps were created and corrected to guarantee that our work was comprehensive, consistent, explicit, reproducible, and of high quality.

### 3.3.4 *Practical Screen*

To select and identify relevant literature we used recommendations provided by Webster and Watson (2002), as well as certain criteria.

- **Content:** we focused on the topics of the literature rather than restricting to one research methodology, one set of journals, or geographic region.
- **Publication language:** We did, however, restrict ourselves to English-written literature.
- **Setting:** only considering studies conducted by/for the IT industry.
- **Date of publication:** We only included papers created after 2016 in our quest for the latest technical solutions and organizational strategies unless they contained information that was still relevant today. Other sorts of articles that were not dependent on improved technical solutions, such as reports or surveys that describe specific technologies or research methodologies, were not subjected to a publication date.

### 3.3.5 *Search for Literature*

Our literature search began with a focus on the Senior Scholars basket of eight publications as well as a few additional cybersecurity-related periodicals. Then, to broaden our search, we used IEEE Xplore, AIS Electronic Library, Google Scholar, ProQuest, and Mendeley, among other search engines and databases. Since we used Mendeley, it also presented us with some personalized recommendations. In the end we also tried the "snowballing" technique, but it did not yield any useful results.

Relevant literature was downloaded and kept in a shared folder in Mendeley's cloud storage during the search, ensuring that duplicate articles were not preserved while also allowing us to share comments and keep track of citations.

When looking for relevant literature, the following search keywords were used:

- Intrusion Detection System
- Artificial Intelligence in Intrusion Detection Systems
- Automation in Intrusion Detection System

- Cybersecurity and Machine learning
- Deep learning Cybersecurity
- Situational Awareness
- Intrusion Detection Situational Awareness
- Security Operation Center
- Automation in Cybersecurity
- Artificial intelligence in Cybersecurity

### ***3.3.6 Appraise Quality***

The search terms and practical screen yielded sixty. These steps do not consider the quality, thus we applied stricter criteria, which led to the exclusion of twenty-six papers that were not relevant after further examination, resulting in thirty-four studies forming the body of our literature review which are listed below. At least one of the requirements listed below had to be met by the literature:

- They had to focus on either IDS, SOC, and/or ML.
- The literature is written in proper English.
- It identifies gaps in existing IDSs.
- It should address a clearly focused question.
- The outcomes must be reliable and applicable to the intended audience.
- Does the literature find and evaluate ML models or frameworks that can be used in conjunction with IDSs?

### ***3.3.7 Extract Data***

We meticulously extracted information that could be used as raw material for the synthesis step from the thirty-four papers that remained and saved the data for each paper using the comment section in Mendeley. An overview of remaining papers that we used to extract data can be found in Table 1 Literature overview shown below. The table sorts the topics in columns and shows every topic that each article covers or mentions in a relevant way. This provided us with a valuable overview of the literature.

Table 1 Literature overview

<i>Article</i>	<i>IDS</i>			<i>SOC</i>			<i>ML</i>		<i>Situational Awareness</i>
	<i>IDS General</i>	<i>NIDS</i>	<i>Anomaly Detection</i>	<i>Triage</i>	<i>Automation</i>	<i>Issues</i>	<i>Deep Learning</i>	<i>Shallow Learning</i>	
Ahmad, A et al., 2021	✓					✓			✓
Ahmad, Z et al., 2021		✓					✓	✓	
Aldweesh et al., 2020							✓		
Apruzzese et al., 2018							✓	✓	
Berman et al., 2019							✓		
Chamkar et al., 2021						✓			
Danquah, 2020				✓	✓				
Dixit & Silakari, 2021							✓		
Fauri et al., 2019	✓		✓				✓	✓	✓
Ferrag et al., 2020							✓		
Gümüşbas et al., 2021								✓	
Gupta et al., 2019					✓	✓	✓		
Karatas et al., 2019							✓		
Kasongo & Sun, 2019							✓		
Kinyua & Awuah, 2021					✓			✓	
Khraisat et al., 2019	✓								
Li et al., 2019	✓						✓		
Liu & Lang, 2019							✓	✓	
Mahdavifar & Ghorbani, 2019							✓		
Martin et al., 2021									✓
Masduki et al., 2017	✓							✓	✓
Mohsienuddin & Lakshmisri, 2018				✓					
Moraboena et al., 2020							✓		
Naseer et al., 2018		✓	✓				✓		
Rahul et al., 2018							✓	✓	

Salitin & Zolait, 2020									✓
Shaukat et al., 2020							✓	✓	
Shone et al., 2018	✓	✓					✓		
Sun et al., 2019	✓								
Vielberth et al., 2020						✓			
Vinayakumar et al., 2019			✓				✓		
Wang, 2018	✓							✓	
Xin et al., 2018							✓	✓	
Zhong et al., 2019				✓	✓				

### ***3.3.8 Synthesis of Studies and Writing the Review***

Following the collection of data, we analyzed, debated, categorized, and compared what we had and constructed the table shown in Table 2. Categorical overview of literature to help us keep track of everything and gain a better understanding by grouping the articles by subjects relevant to our research question. This allowed us to limit down the important topics to those that had already been discussed.

Table 2 Categorical overview of literature

<i>Topic</i>		<i>Description</i>	<i>Articles</i>
<i>IDS</i>	<b>IDS General</b>	Definitions and explanations regarding IDS and the use of the technology. Areas of improvement	(A. Ahmad et al., 2021; Fauri et al., 2019; Khraisat et al., 2019; J. Li et al., 2019; Masduki et al., 2017; Shone et al., 2018; Sun et al., 2019; Wang, 2018)
	<b>NIDS</b>	Definitions and explanations regarding NIDS and the use of the technology. Areas of improvement	(Z. Ahmad et al., 2021; Naseer et al., 2018; Shone et al., 2018)
	<b>Anomaly Detection</b>	Definitions and explanations regarding AIDS and the use of the technology. Areas of improvement	(Fauri et al., 2019; Naseer et al., 2018; Vinayakumar et al., 2019)
<i>SOC</i>	<b>Triage</b>	Explanations of triage in SOC. Issues, and how to improve state of the art. Automation.	(Danquah, 2020; Mohsienuddin Mohammad & Lakshmisri, 2018; Zhong et al., 2019)
	<b>Automation</b>	Automation as the next step in SOC. Too many events to handle by humans.	(Danquah, 2020; Gupta et al., 2019; Kinyua & Awuah, 2021; Zhong et al., 2019)
	<b>Issues</b>	Issues concerning the large number of workloads that a SOC produces, are often used to argue for solutions including the implementation of automation.	(A. Ahmad et al., 2021; Chamkar et al., 2021; Gupta et al., 2019; Vielberth et al., 2020)

<i>ML</i>	<b>Deep Learning</b>	Deep Learning and Deep Neural Networks as a tool to improve accuracy in IDS incident detection.	(Z. Ahmad et al., 2021; Aldweesh et al., 2020; Apruzzese et al., 2018; Berman et al., 2019; Dixit & Silakari, 2021; Fauri et al., 2019; Ferrag et al., 2020; Gupta et al., 2019; Karatas et al., 2019; Kasongo & Sun, 2019; J. Li et al., 2019; Liu & Lang, 2019; Mahdavifar & Ghorbani, 2019; Moraboena et al., 2020; Naseer et al., 2018; Rahul et al., 2018; Shaukat et al., 2020; Shone et al., 2018; Vinayakumar et al., 2019; Xin et al., 2018)
	<b>Shallow Learning</b>	Shallow Learning as a tool to improve accuracy in IDS incident detection.	(Z. Ahmad et al., 2021; Apruzzese et al., 2018; Fauri et al., 2019; Gümüşbaş et al., 2021; Kinyua & Awuah, 2021; Liu & Lang, 2019; Masduki et al., 2017; Rahul et al., 2018; Shaukat et al., 2020; Wang, 2018; Xin et al., 2018)
	<b>UEBA</b>	Behavior analytics used to increase accuracy in IDS incident detection.	(Martín et al., 2021; Salitin & Zolait, 2018)
<i>Situational Awareness</i>		Different solutions to enhance Situational Awareness with intrusion detection.	(A. Ahmad et al., 2021; Fauri et al., 2019; Masduki et al., 2017)

### 3.4 Interview with Domain Expert

When performing the interviews with our domain expert, we sent out a consent form that can be found in Appendix A – Consent Form. This consent form was used to inform the participant that participation would be anonymous, as well as some information regarding data collection and data regulation.

## 4 ARTIFACT

Our artifact’s evolution is described in this section. To implement the offered tactics in our framework, we begin by describing the framework we were inspired by. Then there is a breakdown of how and why we included each of the phases that are illustrated.

### 4.1 Reference Framework

Our framework has been inspired by a framework found in previous literature. During our review we found a framework that detailed an improved SOC automated triage solution. This framework provides a clear overview of how AI and ML fits into an automated IDS detection and response solution.

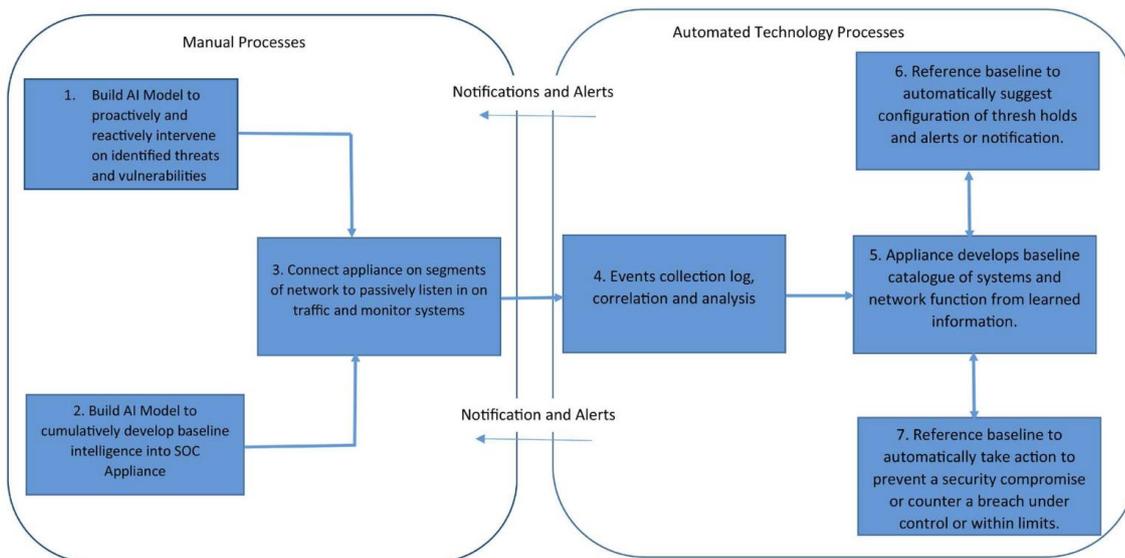


Figure 2 An automated framework for triage, containment, and escalation.  
Source: (Danquah, 2020)

In Figure 2 we can see Danquah’s framework. The framework details a continuous and iterative process that consists of eight required stages, where seven of them are detailed in the figure (Danquah, 2020). On the left side he has detailed manual processes associated with his solution. The “Manual processes” section contains three processes:

1. Build AI model to proactively and reactively intervene on identified threats and vulnerabilities
2. Build AI Model to cumulatively develop baseline intelligence into SOC appliance
3. Connect appliances on segments of the network to passively listen in on traffic and monitor systems.

**Process 1:** The gathered logs from the network's various sources are correlated in this step to determine logic sequences, patterns, and values with the goal of identifying and setting baselines for the monitored network. Baselines are discovered by tracking and comparing events across time to look for consistent sequences of activity. AI is used to identify and describe typical traffic patterns by comparing events from many sources to provide additional information and clarity about patterns on the infrastructure. It can be programmed to continuously learn and react to new evidence while detecting attacks and threats inside the network before they cause a breach.

**Process 2:** This component of the solution makes use of self-learning AI algorithms based on an established baseline of users, devices, systems, and networks within an organization. It either warns IT personnel of potential compliance breaches and threats, or it proceeds to correct a detected breach where the solution can be automated.

**Process 3:** The appliance must be connected to the network in strategic locations. Logging is required at the network's core, distribution, and access levels. Logging is accomplished by passively listening in on traffic to successfully baseline users, devices, and networks inside an organization. The connection is made at several network locations for the purpose of log correlation in order to discover logical sequences.

In the second part of the solution, the author describes the automatic part of the appliance. This section has four processes:

4. Events collection log, correlation, and analysis
5. Appliance develops baseline catalog of systems and network functions from learned information
6. Reference baseline to automatically suggest configuration of thresholds and alerts or notifications
7. Reference baseline to automatically take action to prevent a security compromise or counter a breach under control or within limits.

**Process 4:** Correlating gathered logs from the network's multiple sources is critical to reliability, creating accurate baselines, and identifying potential susceptibility to threats and inherent weaknesses. This is done automatically to find logical sequences, consistent patterns, and values with the goal of setting accurate baselines and avoiding false positive threats and vulnerability reports.

**Process 5:** The appliance creates a starting point for comparisons by collecting and correlating logs from various network sources. This comparison point is not static but learns and adapts to new evidence of the supposed starting point while detecting attacks and threats within the network.

**Process 6:** The set starting point must be constantly referred to in order to assess the priority of dealing with incidents based on the severity of the security breach or compromise. The thresholds are proposed automatically based on the Common Vulnerability Scoring System (CVSS). Unless otherwise changed by human interaction, the indicated thresholds become the default configuration.

**Process 7:** Furthermore, unless otherwise altered by human intervention, the system might be programmed to either warn support or IT experts of compliance breaches and potential dangers or proceed to rectify an identified breach when the solution can be automated.

Lastly is the eighth process, which is not detailed in the framework model.

**Process 8:** Technical support employees are notified and alerted by customizing the prefiltering log events into critical, relevant, and meaningful alerts. Based on configuration, IT experts may be contacted to either address a breach or be notified of a potential breach.

## 4.2 Our Proposed Improvements

Based on the findings of other research as well as our own brainstorming, we discovered that Danquah's approach has some room for improvement.

### 4.2.1 Artificial Intelligence

Danquah recommends using two different AI models for his system: one for detection and one for response. This effectively means that, in an ideal approach, each of his models must learn about every sort of attack in order to be able to recognize and respond to the numerous types of attacks that we are aware of.

During our research we found that there is a severe limitation to try and utilize one algorithm to cover a broad area. No one solution can know everything, especially considering that the broader level of knowledge you want an AI model to learn, the more inefficient it becomes at accurately detecting events. It is very beneficial for both performance and accuracy to have numerous ML models that specialize in a given area. Moreover, we found that combining multiple ML algorithms will provide higher predicted performance than any of the individual learning algorithms (Khraisat et al., 2019). There exist numerous ensemble

methods and Boosting, Bagging, and Stacking are some examples that might be used.

Therefore, we propose to not add one model for detection and one for response, but instead create several ML modules for detection and maybe also response. This is to ensure that each model can focus on one aspect of intrusion detection or response in order to become more effective at this.

In addition, we propose the combined use of DL and SL. SL models can be used in the beginning since they are faster to train and optimize and require less data to train, so they can be utilized while the DL models are being trained. Moreover, because they consume fewer resources and have a shorter running time for both training and testing, they should be selected to handle the “easier” tasks such as detecting well known types of events. This would ideally serve to save resources for the system and keep costs down in the long run, while providing a more effective and robust detection solution. This same principle would be applied to the response model. This would allow us to prioritize resources to the models that need it. Shallow algorithms are also easier to understand because they follow precise principles that explain why decisions are made, and they should be prioritized wherever possible. In comparison to DL models, the random forests technique can do extremely well on a task like zero-day vulnerability detection and serves as an excellent reminder to be cautious when selecting an algorithm to use (Abri et al., 2019; Khraisat et al., 2019). SL models could also handle the “easier” tasks such as detecting well known types of events. This would ideally serve to save resources for the system and keep costs down in the long run, while providing a more effective and robust detection solution. This same principle would be applied to the response model.

When it comes to enhancing the identification of unknown and unanticipated attacks, we recommend the use of DL, as this is the most recommended method for this. Several papers recommend utilizing multiple Deep Neural Network (DNN) models to construct a versatile and effective IDS that can identify and classify unanticipated and surprising cyberattacks (Z. Ahmad et al., 2021; Berman et al., 2019; Dixit & Silakari, 2021; Gupta et al., 2019; Kasongo & Sun, 2019; Liu & Lang, 2019; Shone et al., 2018; Vinayakumar et al., 2019; Wang, 2018). DL should not be utilized in every area. Also, to reduce the intensive processing we recommend that DL should be utilized within areas of high complexity, such as non-linear hypotheses as well as in domains with large-scale data, such as an IDS (Mahdavifar & Ghorbani, 2019). Using a hybrid method by utilizing both SL and DL models within IDS has gotten less attention but is extremely promising, which is why we recommend the use of this in our framework (Xin et al., 2018).

### **4.2.2 Appliance Connection to Network**

In Danquah's model he proposes to let the solution passively listen to the network. This is something that is important to cover *broadly*. However, soon a new version of the TLS transmission encryption will be released, and this causes issues with decryption. In TLS 1.3 we can no longer see basic information about the traffic without decrypting everything, this can cause privacy issues with regards to legality, especially in Norway. Which was an aspect pointed out by the domain expert in Activity 5:

*When listening to the network you should consider the changes that will come with TLS 1.3.*

Thus, a Host based Intrusion Detection System (HIDS) approach is desired. With HIDS you apply the system directly on systems and devices, allowing it to see the decrypted information it needs to assess the risk and purpose of traffic without decrypting the sensitive data.

Therefore, we propose that as well as a network solution to monitor broadly in general, to also apply this automated solution to every value chain that is critical to the organization.

### **4.2.3 Event Log Handling**

Danquah only specifies the collection of events on the network in the logs, however, metrics such as memory usage of systems, CPU usage and others can also provide valuable intel pointing to an attack or event. Therefore, we also include such metrics into the log handling and correlation process.

### **4.2.4 Situational Awareness**

During our research, we discovered that we needed to focus on SA to help incident responders resolve the asymmetry between cyber attackers' attempts to penetrate organizational security and incident responders' attempts to defend the organization from attack while navigating the intricacies of incident response. In a dynamic and complex environment, keeping track of all information events, including the organization's own actions, can be difficult. That is why it is critical to ensure that any potentially relevant information is presented precisely, with the appropriate abstraction and accumulation to compensate for human perception and cognition limitations.

According to our findings, the initial stage of any framework attempting to cover the broad area of security should include a phase that ensures that, the knowledge of knowing what is going on around you, is utilized to determine which events in those surroundings are significant. Since Danquah's approach does not explicitly address SA, we recommend including a phase in the manual process to track SA and ensure that knowledge and tactics are updated throughout the iterative process.

We also visualize how and where the automated process utilizes the three levels of technical SA, namely:

1. **Perception**

The initial level, perception, requires organizations to gather information such alerts and raw data that will be utilized to build the incident-related operational picture later.

2. **Comprehension**

The second level, stands for the comprehension of events by correlating the context received from a tactical viewpoint and necessitates the creation of an operational picture by exposing the impact to its operations by correlating the context obtained from a tactical perspective as well as an understanding of its relevance in terms of cybersecurity goals.

3. **Projection**

The final step (level) is projection, which demands organizations to extrapolate continuously from the operational picture to produce alternative future scenarios of possible threat actors and recognizing patterns in their behavior.

By ensuring that all three levels are covered, the information required to gain SA is delivered to analysts for further examination in the previously mentioned manual phase.

### **4.3 Our framework**

This section covers our improved framework based on Danquah's framework. We show a model of our framework, and explain each step-in detail, as well as explain the differences between our and Danquah's framework.

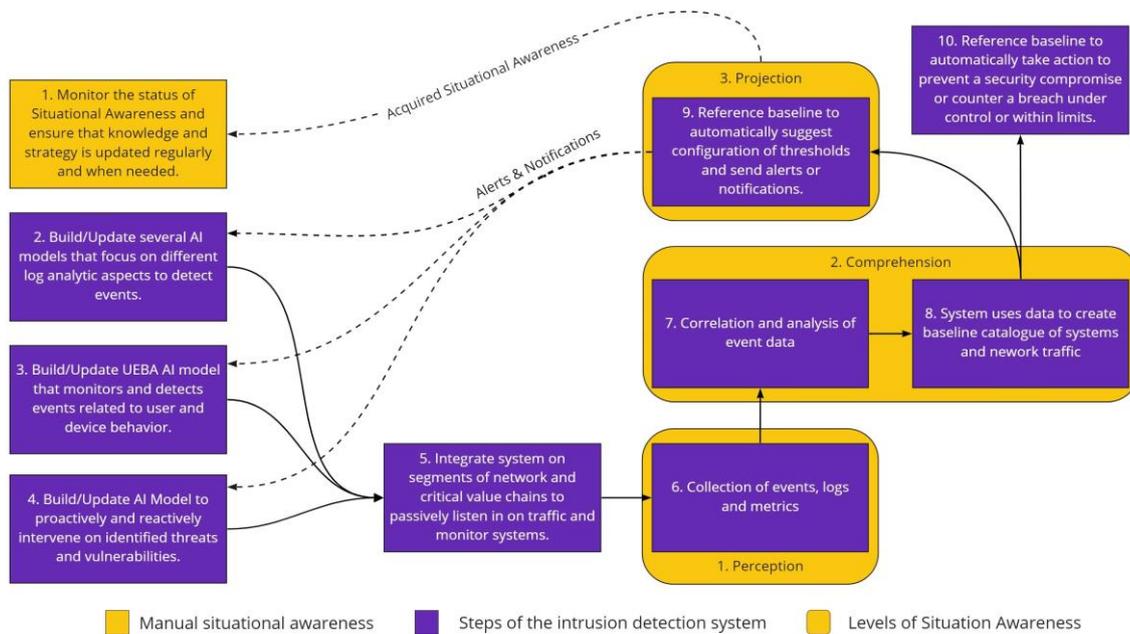


Figure 3 Improved framework for automated security using artificial intelligence and situational awareness.

Figure 3 above shows our framework model. We have added a few steps as well as outlining the role and function of SA in our framework. In this framework, we have tried to address all of the problems we highlighted in the previous section. We accomplish this by adding the use of several AI models that each are specialized in one area of detection. We also extend the collection of data to encompass metrics from system hardware, as this can provide valuable information when scanning for intrusion. Finally, the outer shells that separated the manual and automated process were removed to illustrate that all steps are part of continuous and integrated process.

#### 4.3.1 The detailed phases

This section covers each phase of our framework in detail.

##### *Phase 1: Acquiring Situational Awareness*

The primary phase details the manual process of addressing the state of SA. The fundamental goal of this phase is to maintain a consistently high degree of SA across the computer network's complex and dynamic state, which includes all of the many network objects. To secure all valuable assets, the network administrators should make informed decisions to guarantee that potential threats, the impact of an attack, and risk precautions are all recognized. To keep SA updated, network administrators attain new incident response-related suggestions,

which are visualized as mental models, as well as new strategic threat intelligence regarding threat-actors.

#### *Phase 2: AI Models for Baseline*

In this step, the gathered logs from the various sources in the system are correlated to determine logic sequences, patterns, and values with the goal of identifying and setting baselines for the monitored network. The baselines are discovered by tracking and comparing events across time to look for consistent sequences of activity. Different AI models are used to identify and describe typical traffic patterns by comparing events from many sources to provide additional information and clarify about patterns on the infrastructure. They can be programmed to continuously learn and react to new evidence while detecting attacks and threats inside the network before they cause a breach. The use of different AI models with separate specific focus areas helps provide a better accuracy in the pattern detection process. The domain expert expressed concerns regarding the flow of data and the use of resources to accomplish this. He then explained a way to alleviate the issue:

*Regarding the cost of getting the information for the AI models you should consider using a data streaming processor.*

#### *Phase 3: User and Entity Behavior Analytics*

In this step we define a separate AI model that is specifically focused on monitoring users and devices and developing a baseline knowledge on how each user and device operates under normal operation. This model should also include a “weighting” that can be set on specific users and devices that are known to operate in strange patterns that are hard to define. This would allow IT employees to operate normally without the system flagging their activity at every turn. This will open the system up to attack on these users and devices but will be a compromise that is beneficial to take to ensure normal workflows can continue. This was pointed out to us during the interview with the domain expert:

*The problem with finding anomalies in the behavior of IT employees is the fact that they never follow a standard procedure.*

#### *Phase 4: Orchestration*

This component of the solution makes use of self-learning AI algorithms based on an established baseline of users, devices, systems, and networks within an organization. It either warns IT personnel of potential compliance breaches and threats, or it proceeds to correct a detected breach where the solution can be automated.

Again, like in step 2, we employ different, specialized AI models that help ensure better accuracy in the detection and response to different threats.

#### *Phase 5: System integration*

The system will be integrated to strategic network nodes to passively listen to and log traffic and events. Furthermore, we propose connecting the solution to certain devices and systems in order to monitor specific value chains. This will be especially relevant in the future as network encryption soon will be stricter and no longer allow certain metadata to be decrypted without also compromising potentially sensitive data in the traffic. Thus, a HIDS solution will allow the traffic to be decrypted carefully at the device after the traffic is received or before it is sent. As quoted by the DE:

*Making use of HIDS in this stage will make sure the system acquires the whole value chain.*

Clever positioning of the solution on network, devices and value chains allows for the collection of valuable log and metric data.

#### *Phase 6: Log and Metrics Collection (Perception)*

The solution will collect log data from different sources based on its placement on the network. In addition, it should include metrics such as memory usage, CPU usage and other metrics that can indicate unusual usage of resources. A point that was pointed out by the DE.

*Logs normally only refer to text and including metrics will help specify other types of data that should be collected.*

Perception is the first level in generating SA, and it comprises gathering logs and metrics such as behavioral analytics, internally generated phishing warnings, and data from threat intelligence technologies.

#### *Phase 7: Log and Event Correlation and Analysis (Comprehension)*

Correlating logs from numerous sources on the network and devices is crucial for reliability, generating accurate baselines, and recognizing potential vulnerability to threats and inherent flaws. This is done automatically to uncover logical sequences, consistent patterns, and values in order to build correct baselines and eliminate false positive threats and vulnerability reports.

All of this contributes to the creation of an operational picture of the logs and metrics, which is part of the second level of SA, Comprehension, and aids in understanding its relevance in terms of enhancing SA.

*Phase 8: Tactical and Operational overview (Comprehension)*

By collecting and correlating logs from various network sources, the system provides a starting point for comparisons. This comparison point is not static; rather, it learns and adapts to new evidence of the supported starting point while detecting network intrusions and threats. This phase also falls under Comprehension since it provides more context that aids in the process of making sense of the data.

*Phase 9: Autosuggestion (Projection)*

The predefined starting point must be regularly referred to prioritize dealing with occurrences based on the severity of the security breach or compromise. The thresholds are automatically provided based on the Common Vulnerability Scoring System (CVSS). The mentioned thresholds become the default configuration unless otherwise altered by human intervention.

Continuous extrapolation from the operational picture generates alternative future knowledge in this phase, which is part of SA's final level, Projection. The projection generates realistic and likely possibilities, which contribute to network managers' decision-making to update the initial phase of SA acquisition.

*Phase 10: Automatic Response*

Furthermore, unless otherwise changed by human intervention, the system may be designed to either advise support or IT professionals of potential compliance breaches or to proceed to repair an identified breach when the solution may be automated.

## 5 DISCUSSION

In this section of the report, we reflect and discuss our work when developing our framework as well as reflecting on the challenges and limitations that we have faced. We also mention some avenues for future research.

### 5.1 An AI-based framework for enhanced IDS in SOC

The artifact aims to implement several state-of-the-art methods for automated security in the form of IDS and SOC using AI. This was accomplished by performing interviews with a domain expert, as well as a review of previous research on the topics, where we found a framework that implemented some of these aspects and chose to extend this to include some more.

We find that both addressing previous research and domain experts when researching this was very valuable, as the domain expert had some experiences that are often overlooked in scientific research, like how valuable simple information such as hardware metrics can be when monitoring for security events.

#### 5.1.1 *Machine Learning*

##### **Shallow Learning**

The biggest benefit of utilizing ML is the ability to automatically detect unknown attacks, which is why several papers mention that using SL models can be very useful since they are easier to design and construct compared to DL models (Liu & Lang, 2019; Shaukat et al., 2020; Xin et al., 2018). In order to address this, the framework shows that the use of several AI models that specified to the situation or tasks they will perform is critical.

##### **Deep Learning**

Several reports suggest the use of various Deep Neural Network (DNN) models to develop a flexible and effective IDS to detect and classify unforeseen and unpredictable cyberattacks (Z. Ahmad et al., 2021; Berman et al., 2019; Dixit & Silakari, 2021; Gupta et al., 2019; Kasongo & Sun, 2019; Liu & Lang, 2019; Shone et al., 2018; Vinayakumar et al., 2019; Wang, 2018).

The framework addresses this in the same way as SL, by proposing the use of several different AI models in the solution.

### ***5.1.2 Situational Awareness***

According to the literature study of A. Ahmad et al. (2021), SA is a fundamental attribute of organizational incident response, and they describe SA from a technology standpoint as a process of gathering relevant and valuable data, fusing essential aspects of the data together, and obtaining insights from the merged data.

Fauri et al. (2019) propose a role-based system that monitors BACnet building automation networks and uses SA and intrusion detection to discover devices, classify them according to functional tasks, and detect deviations from the assigned roles. They claim that implementing SA and intrusion detection improves the understandability of alerts and the system's adaptability by observing, parsing, and interpreting network messages and extracting useful information about devices by building a network map to provide operators with details about their system, making it easier to detect attacks.

The framework includes SA in a couple of different ways. It addresses the manual process of addressing the state of SA by having the network administrators make informed decisions about potential threats, the impact of attacks and risk precautions, based on feedback from the solution in order to secure all valuable assets.

In addition, it addresses the three levels of SA, perception, comprehension, and projection. It addresses the first step, perception, by gathering logs and metrics such as behavioral analytics, internally generated phishing warnings and data from threat intelligence technologies. The second step, comprehension, is addressed by correlating logs from the various sources and devices in order to improve reliability, generate accurate baselines and recognize potential vulnerability to threats and inherent flaws. In addition, it compares logs and events to the catalogue of previous events. This also provides further context that aids in the process of making sense of the data. The third step, projection, is achieved by continuously extrapolating data and defining new thresholds based on the CVSS standards.

### ***5.1.3 Security Operation Center***

The previous literature identified several issues with conventional SOCs. SOCs are normally led by a security specialist who manually analyzes the alerts to conclude whether or not they truly are malicious and as many of the papers describe, the challenge with analyzing the alerts a SOC produces is often the sheer number of false positives and false-negatives, making it hard to identify threats (Chamkar et al., 2021; Gupta et al., 2019; Lin T et al., 2018; Vielberth et al., 2020).

The framework addresses this by aiming to include highly specialized AI models and using security analysts to train these models and incorporating self-

learning AI as well. Having several specialized AI models will increase the accuracy greatly (Moraboena et al., 2020).

#### **5.1.4 *Intrusion Detection System***

A paper by Sun et al. (2019) write about incident prediction, as opposed to incident detection. They describe it as a system being able to proactively act against unknown threats while providing knowledge on how to improve security instead of just detecting and mitigating threats.

The framework addresses this through its iterative process of continually comparing current and past events and thus, continuously developing new baselines and thresholds as well as evolving its AI models to be more accurate over time.

## **5.2 Challenges & Limitations**

During our work, we had to alter and modify our scope and focus several times. This was due to us not having enough knowledge about the topic in advance, and after some work realized that some of our ideas were not realistic to accomplish in the time-period of a master's thesis.

Our main idea initially was to help develop a ML module to assist in anomaly detection along with Microsoft Sentinel. This is something that initially seemed like a doable task, but with no real knowledge and experience about automated security systems or ML, we realized that this was not feasible.

Our research on how AI and SA can be integrated successfully to improve the performance of IDSs is not without limits. Due to time constraints, we were only able to complete one cycle of Design Science Research. It would have been ideal to improve the artifact by completing several more iterations of the cycle, but more research and development is expected. Furthermore, obtaining feedback from a variety of security specialists would be preferable, as our research has only been evaluated by one.

## **5.3 Future research**

We could have gained a more comprehensive understanding of SA and investigated the interactions between cybersecurity and non-cybersecurity employees at both the strategic and operational levels during the incident response

if we had more time to devote to our investigation. If we had more time, we could have expanded the scope of our investigation to include non-IT personnel.

We expect that the framework produced in this thesis will provide a solid platform for future research on this subject. A system prototype utilizing our framework could provide significant input and aid in proving or disproving the viability of such a solution. In the subject of cybersecurity, a stricter emphasis on SA and elements such as mindfulness and mindlessness in automated systems is both fascinating and relevant.

A key element that should be tested with SOC personnel is the requirement to compensate for limits in human perception and cognition.

Another interesting practical application for our framework could be in the field of Cyber Threat Intelligence (CTI) capabilities. A paper by Amato et al. (2021) states that while digital transformation and the steadily growing info-sharing networks have been providing a critical amount of data that is suitable for CTI analysis, most cyber intelligence units are struggling with handling the vast quantity and variety of collectable data. Our framework, if implemented correctly, could prove a valuable asset in handling this dataflow and also provide the cyber intelligence unit with metrics and metadata that is very valuable for their work.

## 6 CONCLUSION

The most recent findings indicate that cyber-attacks show no sign of slowing down, which is why it is more important than ever to place an emphasis on the development of more robust automated security solutions. As a consequence of this, sophisticated IDS that are able to recognize modern malware are becoming an increasingly crucial component in the process of securing computer systems. In order to design and create such IDSs, one must have a comprehensive grasp of the benefits and drawbacks of existing IDS solutions. In addition, by combining tactical and operational perspectives, a comprehensive image of the dynamic and complex threat environment must be provided to aid organizations in building SA for incident response.

This thesis adds to the theory in a number of ways. To begin, we propose combining several ML algorithms to provide higher predicted performance. Having to rely on a single algorithm to detect anomalies would drastically limit performance and accuracy. We also propose that each algorithm whether combined or singular, should cover a specific area to optimize performance and accuracy ever more. Including a HIDS as an addition to a network solution for general monitoring, allowing it to access the encrypted information it needs to assess the risk and purpose of traffic without decrypting the sensitive data.

Moreover, we add to the body of knowledge on cybersecurity incident response by creating a framework that illustrates the role of management practice in building SA of cybersecurity events. SA helps incident responders overcome the asymmetry between cyber attackers' attempts to infiltrate corporate security and responders' attempts to safeguard the organization while navigating incident response. Keeping track of information events, including the organization's actions, can be difficult in a dynamic and complicated environment. It is important to abstract and accumulate potentially relevant information to compensate for human perception and cognition limitations.

According to our results, the first stage of any security architecture should use the information of what is happening around you to determine which occurrences are relevant.

## REFERENCES

- Abri, F., Siami-Namini, S., Khanghah, M. A., Soltani, F. M., & Namin, A. S. (2019). *Can Machine/Deep Learning Classifiers Detect Zero-Day Malware with High Accuracy?*
- Ahmad, A., Maynard, S. B., Desouza, K. C., Kotsias, J., Whitty, M. T., & Baskerville, R. L. (2021). How can organizations develop situation awareness for incident response: A case study of management practice. *Computers and Security, 101*. <https://doi.org/10.1016/j.cose.2020.102122>
- Ahmad, Z., Shahid Khan, A., Wai Shiang, C., Abdullah, J., & Ahmad, F. (2021). Network intrusion detection system: A systematic study of machine learning and deep learning approaches. *Transactions on Emerging Telecommunications Technologies, 32*(1). <https://doi.org/10.1002/ett.4150>
- Ahmim, A., Maglaras, L., Ferrag, M. A., Derdour, M., & Janicke, H. (2019). A novel hierarchical intrusion detection system based on decision tree and rules-based models. *Proceedings - 15th Annual International Conference on Distributed Computing in Sensor Systems, DCOSS 2019, 228–233*. <https://doi.org/10.1109/DCOSS.2019.00059>
- Aldweesh, A., Derhab, A., & Emam, A. Z. (2020). Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues. *Knowledge-Based Systems, 189*. <https://doi.org/10.1016/j.knosys.2019.105124>
- Almseidin, M., Alzubi, M., Kovacs, S., & Alkasassbeh, M. (2017). *Evaluation of Machine Learning Algorithms for Intrusion Detection System*.
- Amato, G., Ciccarone, S., Digregorio, P., & Natalucci, G. (2021). A Service Architecture for an Enhanced Cyber Threat Intelligence Capability. In A. Armando & M. Colajanni (Eds.), *Proceedings of the Italian Conference on Cybersecurity, ITASEC 2021, All Digital Event, April 7-9, 2021* (Vol. 2940, pp. 436–446). CEUR-WS.org. <http://ceur-ws.org/Vol-2940/paper37.pdf>
- Apruzzese, G., Colajanni, M., Ferretti, L., Guido, A., & Marchetti, M. (2018). On the effectiveness of machine and deep learning for cyber security. *International Conference on Cyber Conflict, CYCON, 2018-May, 371–389*. <https://doi.org/10.23919/CYCON.2018.8405026>
- Bace, R., & Mell, P. (2001). *NIST Special Publication on Intrusion Detection Systems*. *Intrusion Detection Systems*.

- Berman, D. S., Buczak, A. L., Chavis, J. S., & Corbett, C. L. (2019). A survey of deep learning methods for cyber security. In *Information (Switzerland)* (Vol. 10, Issue 4). MDPI AG. <https://doi.org/10.3390/info10040122>
- Brewer, R. (2019). Could SOAR save skills-short SOCs? *Computer Fraud & Security*, 2019(10), 8–11. [https://doi.org/10.1016/S1361-3723\(19\)30106-X](https://doi.org/10.1016/S1361-3723(19)30106-X)
- Bringhenti, D., Marchetto, G., Sisto, R., Valenza, F., & Yusupov, J. (2019). Towards a fully automated and optimized network security functions orchestration. *2019 4th International Conference on Computing, Communications and Security, ICCCS 2019*. <https://doi.org/10.1109/CCCS.2019.8888130>
- Chamkar, S. A., Maleh, Y., & Gherabi, N. (2021). THE HUMAN FACTOR CAPABILITIES IN SECURITY OPERATION CENTER (SOC). *EDPACS*. <https://doi.org/10.1080/07366981.2021.1977026>
- Danquah, P. (2020). Security Operations Center: A Framework for Automated Triage, Containment and Escalation. *Journal of Information Security*, 11(04), 225–240. <https://doi.org/10.4236/jis.2020.114015>
- de Reuver, M., Sørensen, C., & Basole, R. C. (2017). *The digital platform: a research agenda*. <https://doi.org/10.1057/s41265>
- Deng, Q., & Ji, S. (2018). A Review of Design Science Research in Information Systems: Concept, Process, Outcome, and Evaluation. *Pacific Asia Journal of the Association for Information Systems*, 1–36. <https://doi.org/10.17705/1pais.10101>
- Dixit, P., & Silakari, S. (2021). Deep Learning Algorithms for Cybersecurity Applications: A Technological and Status Review. In *Computer Science Review* (Vol. 39). Elsevier Ireland Ltd. <https://doi.org/10.1016/j.cosrev.2020.100317>
- Enisa. (2021). *ENISA THREAT LANDSCAPE 2021*. <https://doi.org/10.2824/324797>
- Fauri, D., Kapsalakis, M., Ricardo dos Santos, D., Costante, E., den Hartog, J., & Etalle, S. (2019). *Role Inference + Anomaly Detection = Situational Awareness in BACnet Networks*.
- Ferrag, M. A., Maglaras, L., Moschoyiannis, S., & Janicke, H. (2020). Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. *Journal of Information Security and Applications*, 50. <https://doi.org/10.1016/j.jisa.2019.102419>
- Gümüşbaş, D., Yıldırım, T., Genovese, A., & Scotti, F. (2021). A comprehensive survey of databases and deep learning methods for cybersecurity and intrusion detection systems. In *IEEE Systems Journal* (Vol. 15, Issue 2, pp.

- 1717–1731). Institute of Electrical and Electronics Engineers Inc.  
<https://doi.org/10.1109/JSYST.2020.2992966>
- Gupta, N., Traore, I., & de Faria Quinan, P. M. (2019). *Automated Event Prioritization for Security Operation Center using Deep Learning*.
- Karatas, G., Demir, O., & Sahingoz, O. K. (2019). Deep Learning in Intrusion Detection Systems. *International Congress on Big Data, Deep Learning and Fighting Cyber Terrorism, IBIGDELFT 2018 - Proceedings*, 113–116.  
<https://doi.org/10.1109/IBIGDELFT.2018.8625278>
- Kasongo, S. M., & Sun, Y. (2019). A deep learning method with filter based feature engineering for wireless intrusion detection system. *IEEE Access*, 7, 38597–38607. <https://doi.org/10.1109/ACCESS.2019.2905633>
- Khraisat, A., Gondal, I., Vamplew, P., & Kamruzzaman, J. (2019). Survey of intrusion detection systems: techniques, datasets and challenges. *Cybersecurity*, 2(1). <https://doi.org/10.1186/s42400-019-0038-7>
- Kim, D. (Information technology security consultant), & Solomon, M. (Michael G. ). (2019). *Fundamentals of information systems security*.
- Kinyua, J., & Awuah, L. (2021). Ai/ml in security orchestration, automation and response: Future research directions. *Intelligent Automation and Soft Computing*, 28(2), 527–545. <https://doi.org/10.32604/iasc.2021.016240>
- Li, J., Qu, Y., Chao, F., Shum, H. P. H., Ho, E. S. L., & Yang, L. (2019). Machine learning algorithms for network intrusion detection. In *Intelligent Systems Reference Library* (Vol. 151, pp. 151–179). Springer Science and Business Media Deutschland GmbH. [https://doi.org/10.1007/978-3-319-98842-9\\_6](https://doi.org/10.1007/978-3-319-98842-9_6)
- Li, Y., & Liu, Q. (2021). A comprehensive review study of cyber-attacks and cyber security; Emerging trends and recent developments. *Energy Reports*, 7, 8176–8186. <https://doi.org/10.1016/J.EGYR.2021.08.126>
- Lin T, Zhong C, Yen J, & Liu P. (2018). *Retrieval of Relevant Historical Data Triage Operations in Security Operation Centers* (P. Samarati, I. Ray, & I. Ray, Eds.; Vol. 11170). Springer International Publishing. <https://doi.org/10.1007/978-3-030-04834-1>
- Liu, H., & Lang, B. (2019). Machine learning and deep learning methods for intrusion detection systems: A survey. In *Applied Sciences (Switzerland)* (Vol. 9, Issue 20). MDPI AG. <https://doi.org/10.3390/app9204396>
- Mahdavifar, S., & Ghorbani, A. A. (2019). Application of deep learning to cybersecurity: A survey. *Neurocomputing*, 347, 149–176.  
<https://doi.org/10.1016/j.neucom.2019.02.056>
- Martín, A. G., Beltrán, M., Fernández-Isabel, A., & Martín de Diego, I. (2021). An approach to detect user behaviour anomalies within identity

- federations. *Computers and Security*, 108.  
<https://doi.org/10.1016/j.cose.2021.102356>
- Masduki, B. W., Ramli, K., & Salman, M. (2017). *Leverage Intrusion Detection System Framework For Cyber Situational Awareness System*.
- McAfee. (n.d.). *What is a Security Operation System (SOC)?* . Retrieved February 5, 2022, from <https://www.mcafee.com/enterprise/en-us/security-awareness/operations/what-is-soc.html>
- Mishra, P., Varadharajan, V., Tupakula, U., & Pilli, E. S. (2019). A detailed investigation and analysis of using machine learning techniques for intrusion detection. *IEEE Communications Surveys and Tutorials*, 21(1), 686–728. <https://doi.org/10.1109/COMST.2018.2847722>
- MITRE. (2022). *Situation awareness*.  
<https://www.mitre.org/capabilities/cybersecurity/situation-awareness>
- Mohsienuddin Mohammad, S., & Lakshmisri, S. (2018). Security automation in Information technology. In *International Journal of Creative Research Thoughts* (Vol. 6, Issue 2).
- Moraboen, S., Ketepalli, G., & Ragam, P. (2020). A deep learning approach to network intrusion detection using deep autoencoder. *Revue d'Intelligence Artificielle*, 34(4), 457–463. <https://doi.org/10.18280/ria.340410>
- Naseer, S., Saleem, Y., Khalid, S., Bashir, M. K., Han, J., Iqbal, M. M., & Han, K. (2018). Enhanced network anomaly detection based on deep neural networks. *IEEE Access*, 6, 48231–48246.  
<https://doi.org/10.1109/ACCESS.2018.2863036>
- Okoli, C., & Schabram, K. (2012). A Guide to Conducting a Systematic Literature Review of Information Systems Research. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1954824>
- Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3), 45–77.  
<https://doi.org/10.2753/MIS0742-1222240302>
- Pirc, J. (2017, July 5). *The Evolution of Intrusion Detection/Prevention: Then, Now and the Future*. Secureworks. <https://www.secureworks.com/blog/the-evolution-of-intrusion-detection-prevention>
- Ponemon Institute. (2020). *Study on Staffing the IT Security Function in the Age of Automation: United States and United Kingdom ! Ponemon Institute© Research Report*.
- Rahul, V. K., Vinayakumar, R., Soman, K., & Poornachandran, P. (2018, October 16). Evaluating Shallow and Deep Neural Networks for Network Intrusion Detection Systems in Cyber Security. *2018 9th International Conference on*

- Computing, Communication and Networking Technologies, ICCCNT 2018*.  
<https://doi.org/10.1109/ICCCNT.2018.8494096>
- Salitin, M. A., & Zolait, A. H. (2018, November 1). The role of user entity behavior analytics to detect network attacks in real time. *2018 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies, 3ICT 2018*. <https://doi.org/10.1109/3ICT.2018.8855782>
- Şerban, R.-A. (2017). The Impact of Big Data, Sustainability, and Digitalization on Company Performance. *Studies in Business and Economics*, 12(3), 181–189. <https://doi.org/10.1515/sbe-2017-0045>
- Shaukat, K., Luo, S., Varadharajan, V., Hameed, I. A., & Xu, M. (2020). A Survey on Machine Learning Techniques for Cyber Security in the Last Decade. *IEEE Access*, 8, 222310–222354. <https://doi.org/10.1109/ACCESS.2020.3041951>
- Shone, N., Ngoc, T. N., Phai, V. D., & Shi, Q. (2018). A Deep Learning Approach to Network Intrusion Detection. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2(1), 41–50. <https://doi.org/10.1109/TETCI.2017.2772792>
- Sun, N., Zhang, J., Rimba, P., Gao, S., Zhang, L. Y., & Xiang, Y. (2019). Data-Driven Cybersecurity Incident Prediction: A Survey. *IEEE Communications Surveys and Tutorials*, 21(2), 1744–1772. <https://doi.org/10.1109/COMST.2018.2885561>
- Vast, R., Sawant, S., Thorbole, A., & Badgujar, V. (2021). Artificial Intelligence based Security Orchestration, Automation and Response System. *2021 6th International Conference for Convergence in Technology, I2CT 2021*. <https://doi.org/10.1109/I2CT51068.2021.9418109>
- Vielberth, M., Bohm, F., Fichtinger, I., & Pernul, G. (2020). Security Operations Center: A Systematic Study and Open Challenges. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2020.3045514>
- Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., Al-Nemrat, A., & Venkatraman, S. (2019). Deep Learning Approach for Intelligent Intrusion Detection System. *IEEE Access*, 7, 41525–41550. <https://doi.org/10.1109/ACCESS.2019.2895334>
- Wang, Z. (2018). Deep Learning-Based Intrusion Detection with Adversaries. *IEEE Access*, 6, 38367–38384. <https://doi.org/10.1109/ACCESS.2018.2854599>
- Webster, J., & Watson, R. T. (2002). ANALYZING THE PAST TO PREPARE FOR THE FUTURE: WRITING A LITERATURE REVIEW. In *MIS Quarterly* (Vol. 26, Issue 2). <http://www.misq.org/misreview/announce.html>

- Xin, Y., Kong, L., Liu, Z., Chen, Y., Li, Y., Zhu, H., Gao, M., Hou, H., & Wang, C. (2018). Machine Learning and Deep Learning Methods for Cybersecurity. *IEEE Access*, 6, 35365–35381. <https://doi.org/10.1109/ACCESS.2018.2836950>
- Zeadally, S., Adi, E., Baig, Z., & Khan, I. A. (2020). Harnessing artificial intelligence capabilities to improve cybersecurity. *IEEE Access*, 8, 23817–23837. <https://doi.org/10.1109/ACCESS.2020.2968045>
- Zhong, C., Yen, J., Liu, P., & Erbacher, R. F. (2019). Learning from experts' experience: Toward automated cyber security data triage. *IEEE Systems Journal*, 13(1), 603–614. <https://doi.org/10.1109/JSYST.2018.2828832>

## APPENDIX

### Appendix A – Consent Form

#### Vil du delta i forskningsprosjektet

#### *”A framework for improving intrusion detection systems by combining artificial intelligence and situational awareness”?*

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å komme med forslag til forbedring i automatiserte sikkerhetsløsninger. I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

#### **Formål**

Vi vil se på eksisterende løsninger og tidligere forskning på systemer som automatiserer sikkerhet, finne svakheter her, og prøve å komme med forslag til forbedring.

Vi antar at fokus blir på deteksjon av hendelser i systemene, men begrenser oss ikke til det i utgangspunktet.

Dette er forskning i sammenheng med masteroppgave i cybersikkerhet.

#### **Hvem er ansvarlig for forskningsprosjektet?**

Institutt for Informasjonssystemer ved Universitetet i Agder er ansvarlig for prosjektet.

#### **Hvorfor får du spørsmål om å delta?**

Vi vil prate med eksperter innenfor cybersikkerhet.

#### **Hva innebærer det for deg å delta?**

Hvis du velger å delta i prosjektet, innebærer det at du deltar i et intervju. Det vil ta deg ca. 1 time. Intervjuet inneholder spørsmål om verktøy og systemer som brukes hos bedriften i dag, eventuelle styrker og svakheter ved slike systemer, og tanker om forbedringer innenfor dette. Intervjuet vil bli tatt opp via diktafon, og skal transkriberes. Sluttresultatet vil være anonymt.

#### **Det er frivillig å delta**

Det er frivillig å delta i prosjektet. Hvis du velger å delta, kan du når som helst trekke samtykket tilbake uten å oppgi noen grunn. Alle dine personopplysninger vil da bli slettet. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg.

**Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger**

Vi vil bare bruke opplysningene om deg til formålene vi har fortalt om i dette skrivet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket.

- Studentgruppen og vår veileder for masteroppgaven vil ha tilgang til opplysningene.
- Vi vil ikke skrive ned navn eller annen informasjon som kan identifisere deg direkte

**Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?**

Opplysningene anonymiseres når prosjektet avsluttes/oppgaven er godkjent, noe som etter planen er juni 2022. Etter dette vil alle opptak være slettet.

**Dine rettigheter**

Så lenge du kan identifiseres i datamaterialet, har du rett til:

- innsyn i hvilke personopplysninger som er registrert om deg, og å få utlevert en kopi av opplysningene,
- å få rettet personopplysninger om deg,
- å få slettet personopplysninger om deg, og
- å sende klage til Datatilsynet om behandlingen av dine personopplysninger.

**Hva gir oss rett til å behandle personopplysninger om deg?**

Vi behandler opplysninger om deg basert på ditt samtykke.

På oppdrag fra Universitetet i Agder har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

**Hvor kan jeg finne ut mer?**

Hvis du har spørsmål til studien, eller ønsker å benytte deg av dine rettigheter, ta kontakt med:

- Universitetet i Agder ved veileder Paolo Spagnoletti.
- Universitetet i Agder ved student Daniel Lindemann.
- Universitetet i Agder ved student Yaguel van der Meij.
- Vårt personvernombud: Trond Hauso – [personvernombudet@uia.no](mailto:personvernombudet@uia.no)

Hvis du har spørsmål knyttet til NSD sin vurdering av prosjektet, kan du ta kontakt med:

- NSD – Norsk senter for forskningsdata AS på epost ([personvertjenester@nsd.no](mailto:personvertjenester@nsd.no)) eller på telefon: 55 58 21 17.

Med vennlig hilsen

Paolo Spagnoletti  
der Meij  
(Forsker/veileder)

Daniel Lindemann

Yaguel van

## Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet «Analyzing problems with automation in network detection systems», og har fått anledning til å stille spørsmål. Jeg samtykker til:

å delta i intervju

Jeg samtykker til at mine opplysninger behandles frem til prosjektet er avsluttet

---

(Signert av prosjektdeltaker, dato)