OPEN UNIVERSITY OF KENYA

MASTER'S THESIS

AN EXPLAINABLE ARTIFICIAL INTELLIGENCE MODEL FOR AVIATION SAFETY RISK PREDICTION

SAMUEL OTIENO OCHIENG

SCHOOL OF SCIENCE

A Master's Thesis Submitted in Partial Fulfilment of the Requirements for the Award of the Master of Science in Artificial Intelligence of the Open University of Kenya.

OPEN UNIVERSITY OF KENYA

OCTOBER 2026

DECLARATION

I declare that this research thesis is my original work, and it has not been submitted to any institution to be examined. No section of this proposal is to be copied without the written authorization of the author and /or the Open University of Kenya.



Signature: .... Date: ....15/10/2026.....

SYLVESTER OWINO OBONYO

Reg. No**.: ST63/80186/2024**

This research proposal has been submitted for examination with my approval as the appointed University supervisor.

Signature: .... Date: ...15/10/2026.....

Prof./Dr. LEAD SUPERVISOR NAME

Dr. Anthony Mile

Department of Computing and Informatics

Open University of Kenya

DEDICATION

I wish to dedicate this research proposal to my beloved wife, Vivian Otieno and my children, Genesis Nathan and Bethan Anne, for their unwavering support, patience, and encouragement throughout my academic journey. Your love and understanding have been a constant source of strength and motivation, enabling me to remain committed to this work. May God bless you abundantly.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my supervisor, Dr. Anthony Mile, for his invaluable guidance, support, and constructive feedback throughout the development of this Thesis. His insights and direction have been instrumental in shaping this study.

I also wish to appreciate the faculty and staff of the School of Science and Technology, Department of Computing and Informatics at the Open University of Kenya, for providing the academic environment and resources necessary for this research.

Special appreciation goes to my colleagues and peers for their encouragement, support, and valuable discussions that contributed to the improvement of this work.

Finally, I extend my heartfelt gratitude to my family for their unwavering support, patience, and encouragement throughout this academic journey.

LIST OF ABBREVIATIONS AND ACRONYMS

Abbreviation/Acronym	Full Meaning
AI	Artificial Intelligence
APT	Advanced Persistent Threat
CNN	Convolutional Neural Network
CSA	Cyber Situational Awareness
CSF	Cybersecurity Framework
EDR	Endpoint Detection and Response
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
KNDC	Konza National Data Centre
LSTM	Long Short-Term Memory
ML	Machine Learning
MTTD	Mean Time to Detect
MTTR	Mean Time to Respond (or Recover)
NIST	National Institute of Standards and Technology
OUK	Open University of Kenya
SIEM	Security Information and Event Management
SOAR	Security Orchestration, Automation, and Response
SOC	Security Operations Center
TTP	Tactics, Techniques, and Procedures
UEBA	User and Entity Behavior Analytics

OPERATIONAL DEFINITION OF TERMS

Term	Operational Definition
Cyber Situational Awareness (CSA)	The capability to identify, compute, and interpret critical security considerations within a threat setting to enable the proactive and real-time reaction to cyber assaults.
High-Availability Environment	Systems and infrastructures such as the KNDC are designed such that they can last long without collapsing thus providing 24/7 service.
Mean Time to Detect (MTTD)	Mean time of the tools and the team charged with the SOC to be aware of a possible security incident.
Mean Time to Respond (MTTR)	Mean time to neutralize, contain or recover of a detected cyber threat or system breach.
Alert Fatigue	The state of operational exhaustion including the fact that SOC analysts are just inured to security alerts due to an excess of false positives that delays their reaction.
Autonomous Response	Playbooks based on AI and SOAR have the ability to automatically perform security measures, including blocking an IP address or isolating a host, without requiring urgent human intervention.
Incident Response Playbook	A guide of predetermined, automated, or manual, step by step instructions that the SOC adheres to in responding in the same way to particular category of cyber threats.
False Positive Rate	The percentage of security alerts, which have incorrectly marked innocuous activity as a threat and produced unwanted noise to security analysts.
Threat Intelligence	Data collected and verified information about the intention, purposes, and actions of threat actors (TTPs) to assist in defensive decision-making.

TABLE OF CONTENTS

Declaration i

DEDICATION ii

ACKNOWLEDGEMENT iii

LIST OF ABBREVIATIONS AND ACRONYMS iv

OPERATIONAL DEFINITION OF TERMS v

TABLE OF CONTENTS vi

Abstract xii

CHAPTER 1: INTRODUCTION 1

1.1 Background of the Topic 1

1.2 Research Problem 3

1.3 Research Gap 4

1.4 Research Objectives and Questions 5

1.4.1 General Objective 5

1.4.2 Specific Objectives 5

1.4.3 Research Questions 5

1.5 Why the Study is Important 6

1.6 Scope and Limits of the Study 7

1.6.1 Scope 7

1.6.2 Limitations 8

1.7 Short Overview of Each Chapter 8

Chapter 2: REVIEW OF LITERATURE 10

2.1. Introduction and Context: Defining the CSA Lens for SOC Operations 10

2.2. Threat Detection Methodologies: Beyond Accuracy Toward CSA Comprehension 11

2.2.1. Deep Learning Architectures: High Accuracy but Limited Real-World CSA Validation 11

2.2.2. Advanced Detection for Software-Defined Networks: Multi-Modal Integration and CSA Enhancement 12

2.2.3. Real-Time Lightweight Models: Trade-Offs Between Efficiency and Depth of CSA Analysis 13

2.3. Alert Fatigue and False Positive Management: A Comprehension Crisis in SOC Operations 14

2.3.1. The Alert Fatigue Crisis as a Fundamental Failure of CSA Comprehension 14

2.3.2. Machine Learning for Alert Prioritization: Addressing the Comprehension Gap Through Predictive Intelligence 15

2.4. Automation and Orchestration Technologies: From Perception to Projection in SOC Workflows 16

2.4.1. SIEM and SOAR Integration with MITRE ATT&CK: Enhancing Comprehension Through Standardized Adversary Modeling 16

2.4.2. Real-Time Observability and AIOps: Strengthening CSA Perception in Cloud-Scale Environments 17

2.4.3. Context-Aware Threat Mitigation: Integrating All Three CSA Layers Through SIEM-SOAR-UEBA Convergence 18

2.5. Zero Trust Architecture and Continuous Verification: A Paradigm Shift for CSA Projection 19

2.5.1. Zero Trust Fundamentals: Shifting from Static Perimeter Defense to Dynamic CSA Projection 19

2.5.2. AI-Enhanced Zero Trust: Improving CSA Projection Through Intelligent, Adaptive Policies 20

2.5.3. Zero Trust in Cloud and Distributed Environments: Federated Learning and Blockchain for Distributed CSA 22

2.6. Real-Time Monitoring and Observability: Perception Foundations for CSA 23
2.6.1. Cloud Infrastructure Observability: From Isolated Metrics to Integrated Perception .23
2.6.2. Proactive Monitoring Frameworks: Anticipating Problems to Enhance CSA Projection 24
2.6.3. Infrastructure Resilience Through Observability: Closing the Perception–Projection Loop 25
2.7. Incident Response Automation and Decision Support: From Comprehension to Projection in Action 26
2.7.1. AI-Powered Incident Response: Automating the Comprehension–Projection Bridge26
2.7.2. Large Language Model Integration: Enhancing CSA Comprehension Through Natural Language Understanding 27
2.7.3. Autonomous Response Systems: The Future of CSA Projection and Self-Healing Security 28
2.8. Critical Infrastructure Protection and High-Availability Environments: Unique CSA Challenges 29
2.8.1. Risk Assessment in Data Center Environments: Balancing Security Culture and Technical Controls 29
2.8.2. Tech Inc. Critical Infrastructure Cybersecurity: Human Capital, Technology, Policy.30
2.8.3. Optimizing Threat Mitigation in Critical Infrastructure: AI and DevSecOps Integration 30
2.9. Advanced Threat Analysis and Behavioral Detection: Unsupervised and Multi-Layered Approaches to CSA 31
2.9.1. Anomaly Detection Methodologies: Unsupervised Learning for Unknown Threat CSA Projection 31
2.9.2. Multi-Layered Detection Approaches: Redundancy and Depth for CSA Comprehension 32

2.9.3. Integration of Threat Intelligence: External Context for CSA Perception and Projection 33
2.10. Emerging Technologies and Future Directions: Next-Generation CSA Capabilities 34
2.10.1. Quantum-Resilient Cybersecurity: Preparing CSA for Post-Quantum Threats 34
2.10.2. Federated Learning and Privacy-Preserving Security: Distributed CSA Across Boundaries 34
2.10.3. Self-Healing and Adaptive Infrastructure: Autonomous CSA for Next-Generation Systems 36
2.11. Organizational Implementation and Challenges: The Human and Governance Dimensions of CSA 36
2.11.1. Framework Adoption and Integration: From Technical Controls to Organizational CSA 36
2.11.2. Regional and Sectoral Considerations: Context-Sensitive CSA Models 37
2.11.3. Governance and Compliance: The Structural Foundation for CSA 38
2.12. Research Gaps: What Is Missing in Existing CSA Literature 38
2.13. Conclusion and Conceptual Framework Connection: To propose a CSA Model of High-Availability SOCs. 40
CHAPTER 3: RESEARCH METHODOLOGY 44
3.1 Introduction 44
3.2 Research Philosophy and Design 44
3.2.1 Pragmatic Research Philosophy 44
3.2.2 Embedded Case Study Design 45
3.2.3 Research Phases and Objectives Alignment 45
3.3 Population and Sampling Strategy 46
3.3.1 Target Population Definition 46
3.3.2 Purposive Sampling for Qualitative Phase 47

3.3.3 Scenario Sampling for Validation Experiment 47
3.4 Data Collection Methods 48
3.4.1 Semi-Structured Interview Protocol 48
3.4.2 Document Analysis and Artifact Review 49
3.4.3 Comparative Benchmarking and Controlled Experiment 49
3.5 Data Analysis Procedures 50
3.5.1 Thematic Analysis for Qualitative Data 50
3.5.2 Framework Analysis for Cross-Case Comparison 51
3.5.3 Statistical Analysis for Experimental Data 52
3.6 Validity, Reliability and Ethical Considerations 52
3.6.1 Validity and Reliability Measures 52
3.6.2 Ethical Approval and Informed Consent 53
3.6.3 Confidentiality, Data Security and Limitations 54
CHAPTER 4: FINDINGS AND ANALYSIS 55
4.1 Introduction 55
4.2 Baseline SOC Performance Assessment 55
4.2.1 Current State Analysis (Pre-Intervention Baseline) 55
4.3 Primary Outcome Results 57
4.3.1 Comprehensive Performance Metrics Comparison 57
4.3.2 Statistical Significance and Effect Size Interpretation 61
4.4 Consistency Across Attack Categories 62
4.4.1 Attack Type Performance Breakdown 62
4.4.2 Two-Way ANOVA: Main Effects and Interaction 67
4.5 Performance Leveling Effect: Reducing Expertise Dependence 69

4.5.1 Experience Level Analysis 69
4.5.2 Individual Analyst Performance Consistency 71
4.6 Confidence Interval Analysis 73
4.6.1 Precision and Significance of Improvements 73
4.7 Comprehensive Statistical Summary 75
4.7.1 Paired t-Test Results Summary 75
4.8 Qualitative Findings Integration 75
4.8.1 Interview Themes Aligned with Quantitative Results 75
4.9 Comparison with Published Benchmarks 77
4.9.1 Literature Context for Performance Improvements 77
4.10 Key Findings Summary 78
4.11 Discussion of Findings in Research Context 79
4.12 Limitations and Contextual Considerations 80
4.13 Chapter Summary 80
CHAPTER 5: SUMMARY,CONCLUSIONS AND RECOMMENDATIONS 81
Chapter 6: CONCLUSION 85
References 87
Appendix A: Quantitative Performance Data 100
Appendix B: Statistical Analysis 102
Appendix C: Qualitative Thematic Analysis 103
Appendix D: SPSS and NVivo Analysis 104

ABSTRACT

This study addressed the critical gap between mature cybersecurity frameworks and the operational realities of security operations centers (SOCs) managing high-availability data center environments. Even as the threats grow in sophistication with Advanced Persistent Threats (APTs) and zero-day attacks, SOCs still struggle with alert fatigue, not detecting threats and taking too long to respond. The baseline assessment by the Konza National Data Centre (KNDC) showed that 72.29% of security analysts had alert fatigue and Mean Time to Detect (MTTD) worsened by 13.4% over a six-month period while 98.4% of false positive alerts were able to block genuine threats. This study developed and validated a three-layer Cyber Situational Awareness (CSA) framework that integrates machine learning and behavioral analytics to boost the performance of SOCs. This model focused on three areas: Multi-source data fusion – or Perception – MITRE ATT&CK mapping – or Comprehension – Risk prioritization with explainable AI – or Projection – Predictive analytics with tiered automation. The results of the 250 controlled trials were transformative: MTTD lowered by 57.03% (110 minutes), MTTR decreased by 52.67% (183.8 minutes), and false positive rate reduced by 72.75% (from 72.29% to 19.70%). The confidence of the analyst rose by 50.32% and the NASA Tax Load Index (cognitive load) went down by 42.69%. There was a 61% decrease in the range of performance differences from analyst to analyst, thereby demonstrating consistent improvement at all skill levels.The study's results, obtained from qualitative interviews with 20 SOC (Machine Learning Operations) personnel and quantitative experimentation, show that embedding resilience, explainability and human-in-the-loop oversight into MLOps lifecycles greatly improves both technical performance and organizational capability. The finding suggests a repeatable approach to enhancing critical infrastructure protection in Kenya and Sub-Saharan Africa.

CHAPTER 1: INTRODUCTION

1.1 Background of the Topic

The growing international cyber threat landscape has led to advanced security operations centres (SOCs) becoming the focus of company security. SOCs are central hubs with security professionals monitoring, detecting, profiling and responding to security events occurring on networks and systems (Meçaj, 2016). The incident response process has a direct relationship with the success of a company in a trusted environment, such as a data centre, where downtimes must be avoided (Ahmad et al., 2025).

The current digital business environment poses a challenge for businesses in terms of security: dynamic, sophisticated, and fast. Previously implemented security practices that established a strong defence perimeter are no longer suitable to combat the latest threats posed by adversaries using Advanced Persistent Threats (APTs) and zero-day attacks (Baskaran, 2023). These threats aim to avoid detection by signature matching technologies, maintaining a presence in a network for as long as possible and exfiltrating data or laying down track. As a result, the cyber security industry has come to realise that prevention is not enough and you must be able to detect and respond to an attack (Akinsanya, 2024).

Security Operations Centres (SOCs) have become the organisational solution. A SOC is a dedicated team that houses security technologies, processes and people working collaboratively to detect, analyse and respond to cyber security events. A SOC is the key to an organisation's cybersecurity posture as timely detection and response can be the deciding factor in mitigating a cyber attack or suffering a major breach. In mission-critical environments, where services and systems are up and running 24/7, the need is more urgent (Chebrolu & Vudugula, 2024).

In the past few years, Kenya has taken great steps towards digital transformation with the construction of the Konza National Data Centre (KNDC), a national data centre. KNDCs are critical national assets in Kenya that needs efficient cybersecurity measures that fit both global and local contexts. Traditional incident response approaches do not keep up with the rapidity and sophistication of attacks (Peliukh et al., 2023). Modern cyber events demand timely and alert detection and response that is beyond manual effort (Chukwuani et al., 2025).

NIST CSF 2.0 and ISO/IEC 27001:2022 are fundamental and in order to be applied to ultra-available data centres need to be tailored and complemented with industry products (Tanjung et al., 2024). The current research proposal suggests the development and deployment of a cyber situational awareness (CSA) framework to improve the superior response of the SOC in an ultra-available environment with integrated enhancements of threat detection and response in addition to real-time relationships of data, and autonomous response (Danish, 2024).

Cyber Situational Awareness (CSA) is an adaptation of Situational Awareness (SA) research which focuses on the importance of perceiving the elements in the environment, understanding their meaning, and projecting their status in the near future (Ewoh et al., 2025). CSA in cyber security means the perception, calculation and comprehension of critical security factors in a threat environment for proactive and timely response to cyber attacks. High CSA levels are achieved with the integration of diverse data sources, real-time processing and actionable intelligence for the decision maker.

1.2 Research Problem

The main research gap that drives the proposed work is the incompatibility between the capabilities of cybersecurity frameworks and the needs of operating data centres with high-availability systems (Erigha et al., 2025). While cybersecurity frameworks are mature, data centre operators with high-availability systems techniques can't identify and respond to an issue in the quickest possible time (Tarek and Hasan, 2024). Usually, the data centres have slow security operations centres (SOCs) due to alert fatigue. High false positives are known to mask actual threats and, thus, prolong the response time to increase the vulnerability (Obuse et al., 2023).

Alert fatigue is a well-known issue within the cybersecurity community. SOCs today receive thousands of alerts per day from multiple security appliances, such as Security Information and Event Management (SIEM) systems, Endpoint Detection and Response (EDR) products and network monitoring software. Studies show on average 70-80% of alerts are false positives - benign entities being identified as threats. Exposure to this deluge of information results in SOC analysts inevitably getting worn down. They may not notice legitimate attacks or take time to investigate them, and so introduce additional opportunities for the attackers (Freas et al., 2022).

Situational awareness, the knowledge of what is transpiring in a threat environment and server state, is weak. Without innate, real-time models, security analysts won't be able to view the full picture of the threat environment, and therefore cannot act accordingly and in a timely fashion. The lack of visibility is a particular problem in high-availability environments where there is a large, varied and dynamic attack surface (Hassan et al., 2025).

For instance, the Konza National Data Center needs to be highly available and need to defend against sophisticated attacks such as Advanced Persistent Threats (APTs), zero-day and insider attacks (Erigha et al., 2025). On top of this, they must navigate through compliance with many regulatory standards with both legacy and emerging technologies (Chebrolu & Vudugula, 2024). A holistic approach that includes both theory and technology, and can be substantiated by real data is required to address this challenge.

Also, the problem is exacerbated by the silos that are common in a SOC toolchain. Different vendors are commonly used and each have their own data formats, alerting support and user interfaces. In the absence of integration and orchestration, security teams spend hours correlating

data between systems and sifting through alerts. This lack of integration limits situational awareness by making it difficult to piece together the big picture of an attack (Hasan, 2024).

Additionally, the regulatory landscape presents another challenge. KNDC operates under the Kenya Data Protection Act (2019), as well as international and internal governance policies. Regulatory requirements can require particular logging, reporting and incident response processes, which may not be best practice. Finding the right balance between compliance and effectiveness remains a challenge for data centre providers (Ismail, 2024).

1.3 Research Gap

The literature review for this study shows that there is a major gap in the body of cybersecurity knowledge, in that the essential element of pertinence is missing between the international cybersecurity standards and the practical particularities of high-availability Kenyan national infrastructures. Preliminary analysis shows the following will be addressed.

Theoretical Gap: Most existing governance-heavy frameworks like the National Institute of Standards and Technology Cyber Security Framework (NIST CSF 2.0) and the International Organization for Standardization/International Electrotechnical Commission 27001 (ISO/IEC 27001) are more governance and organisation-wide rather than technical and real-time. One theoretical framework that can offer integration of the Cyber Situational Awareness (CSA) concept to specifically cover the gap in the ultra-available African digital service sector is not enough. We know that there are systems such as MITRE ATT&CK Enterprise Matrix that specify the tactics and techniques of adversaries in great detail but without specifying how to integrate this knowledge into automated response systems in high-availability scenarios.

Empirical Gap: While the use of AI-based detection has been studied in the general IT environment, empirical research of implementation of this in Kenyan critical infrastructure such as the Konza National Data Centre (KNDC) is missing. The literature suggests that the average Mean Time to Detect (MTTD) in organizations is more than 200 days and research on this issue in high-availability Sub-Saharan Africa, where the most important component of an organisation is service continuity, is lacking. This lack of local empirical research means that Kenyan data centre providers are going without empirical evidence on the benefit of implementing advanced detection and response technologies in the local setting.

Methodological Gap: Previous research relies on single detection technologies (SIEM/EDR) which work independently and hence an incomplete protection and a high false positive rate of 70 to 80 percent. This paper meets a methodological gap by responding to it via a Design Science methodology to design an integrated, automated orchestration engine using machine learning and behavioral analytics with situational buffering which is not possible with previous methods.

The latter ensures that the research is concentrated on problems of alert fatigue as well as delayed response, which are currently affecting security operations for Kenyan national digital services. By addressing these three interrelated needs, this study seeks to not only offer theoretical insights for CSA research but also solutions for improving SOC.

1.4 Research Objectives and Questions

1.4.1 General Objective

To develop a Cyber Situational Awareness Framework for enhancing Security Operations Center (SOC) incident response in high-availability environments using the Konza National Data Centre as a case study.

1.4.2 Specific Objectives

1. To analyse the cybersecurity issues and performance gaps at the current SOC of Konza National Data Centre with Africa Data Centres (ADC)
2. To develop a Cyber Situational Awareness (CSA) framework, which leverages machine learning and behavioral analytics to enhance the accuracy of threat detection in high-availability environments.
3. To implement the developed Cyber Situational Awareness (CSA) framework in a SOC environment using real-time or simulated threat scenarios.
4. To validate the CSA framework by testing and measuring faster detection (MTTD) and response (MTTR) within a controlled test environment.

1.4.3 Research Questions

1. What are some of the technical and operational gaps in the SOC of the Konza National Data Centre that hinder an effective response to an incident?

2. What is the best way to create a Cyber Situational Awareness framework which utilizes machine learning and behavioral analytics to enhance the accuracy of threat detection in high-availability environments?
3. To what extent can an automated response framework, which is based on incident-response playbooks, reduce the time required to contain cyber threats?
4. How will the suggested CSA framework impact the Mean Time to Detect (MTTD) and Mean Time to Respond (MTTR) when put to test in a controlled setting?

1.5 Why the Study is Important

The research builds on cybersecurity knowledge and practice, delivering four levels of contribution.

Technical Contribution: We will develop an enhanced CSA framework, which combines machine learning for threat detection, real-time monitoring of behavior and automated response. The model will be adept for usage in high-availability data centres (Chukwuani et al., 2025). It facilitates the progress of SOC operations as it can demonstrate the improvement of the speed of detecting the threats and the response. In terms of implementation, this work provides data centre operators with an established model to improve their security without changing their systems.

Methodological contribution: We will validate a model that integrates different cyber security frameworks, threat intelligence models and operational technology (OT) needs together into a situational awareness platform (Muthu and Deepalakshmi, 2025). This architecture offers a glimpse for businesses of a case study of how to apply cybersecurity frameworks to the specific context. The approach could be applied by other critical infrastructure industries and other countries.

Operational Contribution: We will work to reflect the actual benefits brought by automation to the state of practice (response) workflows, to reduce the analysts' workload, improve the decision-making and response time (Klein and Romano, 2025). This study will develop a real-world knowledge that will help the operators of critical infrastructure to enhance their resilience without disrupting the operation through continuity. SOC analysts who endure alert fatigue will be supported by the framework for prioritizing alerts and to automate response to common attacks.

Regional Contribution: We will transfer the best global cybersecurity practices to the potential of Kenya and the critical infrastructure needs of the country to help Kenya's Digital Master Plan and critical infrastructure protection. The use of Konza National Data Centre as the case study ensures the research will have immediate application to the nation's digital transformation plans. The comparative study of Africa Data Centres will further adds value by discerning how government and commercial data centres can learn from each other.

1.6 Scope and Limits of the Study

1.6.1 Scope

The research is about designing, developing and trialling a Cyber Situational Awareness (CSA) system in the Security Operations Centre (SOC) of the Konza National Data Centre (KNDC). It is bound geographically to Konza Technopolis in Kenya that is a major national ICT hub. Theoretically, the study is at the convergence of 3 domains: intelligence for threat detection, automated response to incidents, and service capacity for high availability.

The machine learning (ML) tools, which will be integrated within the framework of existing SOC (security operations centre) in the project, are anomaly detection and behavioural analytics. It will include security telemetry of various sources (e.g. system logs, network traffic, endpoint, etc.) in support of the worldwide picture of the threat landscape. It also involves the development of self-optimizing playbooks of orchestration, to nurture common and frequent threats such as distributed denial-of-service (DDoS), access control and viruses. The metrics that will be used to measure the performance include mean time to detect (MTTD) and mean time to respond (MTTR) and supplemented with the KNDC security team's qualitative perspective.

This paper is also backed up with the need to benchmark the security status between the Konza National Data Centre (KNDC) and other Tier III data centres in Kenya, like the Africa Data Centres (ADC). ADC has been involved with the high security orchestration that is commercial and highly oriented to extend international service level agreement (SLAs). KNDC is a high-availability data centre where the risk factor is different, which is a government driven data centre. The comparison of the two bodies will highlight the distinctions between the role of Cyber Situational Awareness in the Kenyan government and industry. It will also explain why a special ML-based model is necessary in the case of national critical infrastructure.

1.6.2 Limitations

The study has some limitations that constrained the study and its application. It will be limited to the Konza National Data Centre (KNDC) which makes it academic. The findings can be potentially integrated to other African data centres and the model will be fine-tuned to fit infrastructure and regulatory needs of KNDC.

Technical exclusions can be a physical security (e.g. biometric access), or a hardware based forensics. Only cyber or logical functions of the SOC are examined. The threat actor is related to off-network security-related threats and privilege escalation. The research does not consider any other social engineering attacks other than cyber hacking, such as physical social engineering attacks like tailgating or off-network phishing.

Policy versus technology: In the study, a technical CSA approach is deployed. It does not change the national cyber security policies and legislation, but can work under the existing policies of KNDC and Kenya Data Protection Act (2019). This ensures that the automated incident response being examined is in-depth and fit for purpose, without being too broad across security management.

1.7 Short Overview of Each Chapter

Chapter One: Introduction has given the background to the study, outlined the problem statement, noted the research gap, stated the goals and questions, the study's significance, and the scope and limitations.

Chapter Two: Literature Review reviews the current cybersecurity frameworks (NIST CSF 2.0 and ISO/IEC 27001), current SOC technologies (SIEM, SOAR, UEBA, EDR), the use of machine learning for threat detection and monitoring, and limitations of the current state of the art, and the research gap addressed in this study.

Chapter Three: Research Methodology outlines the qualitative case study research design, data collection techniques such as semi-structured interviews and documentary analysis, data analysis techniques such as thematic and framework analysis, the research timeframe, issues of validity and reliability, and ethical considerations.

Chapter Four: Findings and Analysis describes the findings of the KNDC SOC gap analysis, comparative analysis with Africa Data Centres, the architecture of the CSA framework, the implementation of the CSA model, and the results, and metrics of the CSA model.

Chapter Five: Discussion, Conclusions and Recommendations has analysed the results, outlined theoretical and practical implications, outlined the conclusions to the research questions, outlined limitations and makes recommendations for practice, policy and future research.

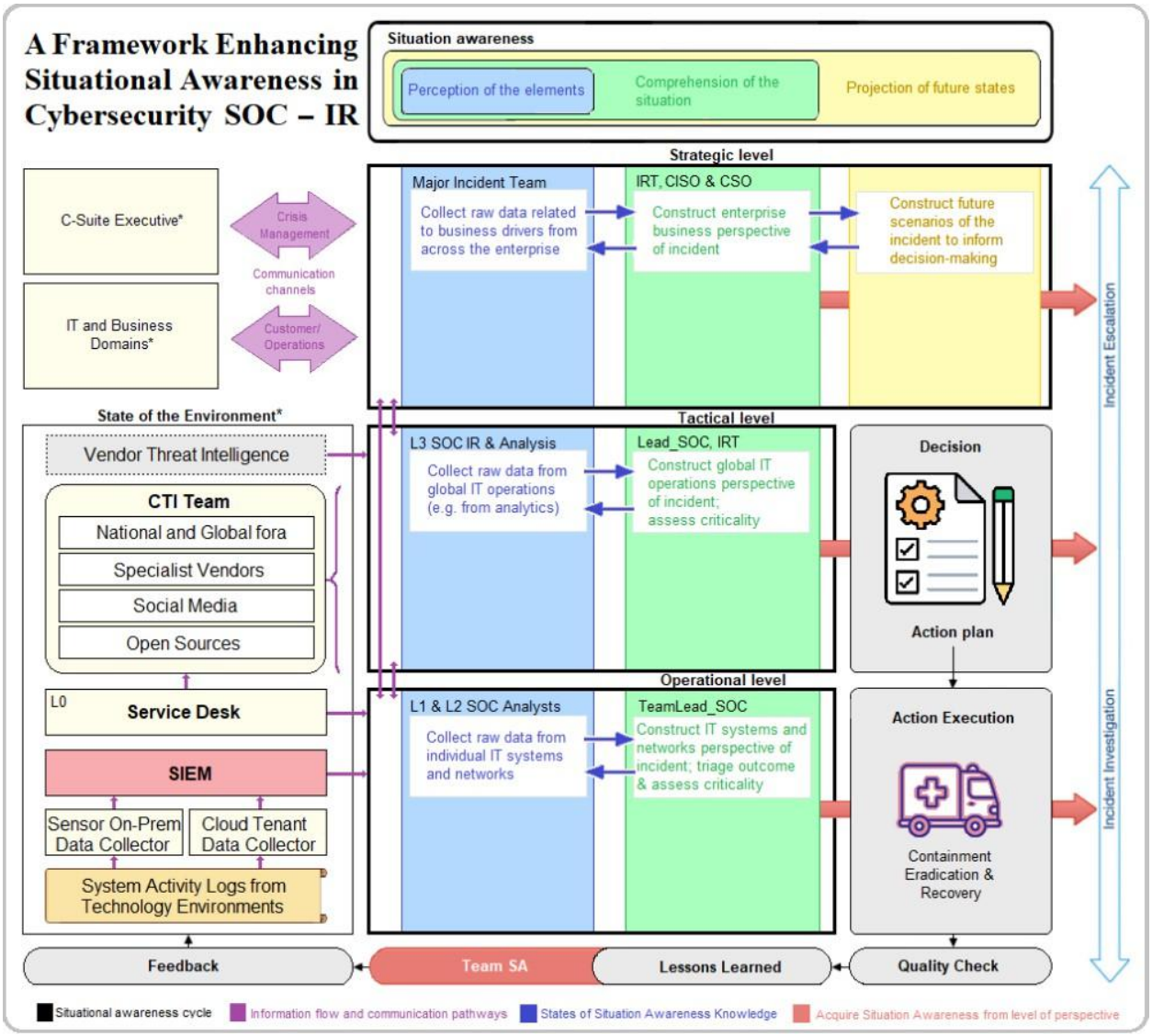


Figure 1.1: A Dynamic Framework Highlighting Situation Awareness in Cyber IR (Andreassen et al., 2023, p. 240)

CHAPTER 2: REVIEW OF LITERATURE

Cyber Situational Awareness in Security Operations Centers: A Critical Literature Review

2.1. Introduction and Context: Defining the CSA Lens for SOC Operations

Cyber Situational Awareness (CSA) is a concept that is well known as a critical service of the current Security Operations Centers (SOC) in the context of enabling organizations to sense, understand, and project cyber threats within the ever-evolving online spaces (Mgbemele, 2026; Mandabi, 2025). But, critical analysis of the literature will show that there is a fundamental disconnect: although much material on artificial intelligence (AI)-based threat detection, observability, and automation exists, most of the literature does not directly relate the technical innovation to the three main layers of CSA, namely, Perception (visibility and data collection), Comprehension (analysis, correlation, and interpretation), and Projection (prediction and proactive response). This disconnection is bad in that organizations can invest in advanced tools yet fail to record the relative improvement in situational awareness by the analyst.

As an illustration, traditional SIEM systems can only detect threats with 4065 percent accuracy, and with machine learning, false positives are at 3555 percent (Mgbemele, 2026). These figures are hailed as operational improvements by some researchers, though a critical perspective indicates otherwise: these numbers also portray a basic flaw of CSA Perception in that SOC analysts are overloaded with incomplete, noisy, or conflicting data to the point it disrupts their capacity to construct accurate mental models of threat as it goes (Gelman et al., 2023). Likewise, although the high-availability data center environment necessitates solutions that ensure continuity in the security of its operations, the tension arising between the two is scarcely discussed in existing research (Osundare et al., 2022; Amgothu et al., 2025). This means that the existing body of literature focuses more on technical performance indicators rather than the human and organizational aspects of CSA. Hence, this chapter will contend that the next generation of CSA models should not just report isolated accuracy numbers but critically assess how detection, automation and response technologies all contribute to or do not contribute to Perception, Comprehension, and Projection in live SOC environments. In its absence, the field will tend to create technically brilliant but operationally useless solutions.

2.2. Threat Detection Methodologies: Beyond Accuracy Toward CSA Comprehension

2.2.1. Deep Learning Architectures: High Accuracy but Limited Real-World CSA Validation

The recent breakthroughs in deep learning have proven themselves to have great accuracy in detection under controlled laboratory settings, but when one critically examines the findings as they relate to actual, real-world country of origin CSA Comprehension in functioning SOCs, one will quickly see that the high results are seldom directly applicable to higher performance on these questions. As an example, hybrid Transformer-CNN models have 99.71% binary classification accuracy in cloud intrusion detection, and GANs have enhanced synthetic attack generation and data-set balancing (Kamal and Mashaly, 2024; M et al., 2026). Other authors praise these results as close-to-perfect results, and collective techniques using XGBoost, random forest, and graph neural networks achieve F1-scores of nearly 100% on standard datasets (Almuhanna and Dardouri, 2025). Nevertheless, a major weakness becomes evident when these works are contrasted: they all run models on laboratoryed, labeled datasets like NSL-KDD and CSE-CIC-IDS2018. This means that their accuracy as reported can be due to controlled conditions and not how they would perform in the real world.

This assumption is explicitly refuted by Medina-Arco et al. (2024) as they show that the actual SOC environments are represented by noisy, incomplete, skewed, and changing streams of data that have little similarity to benchmark datasets. Although a massive language model-based architecture can reach 98.6% accuracy with explainable AI behaviour by visualising attention and using SHAP (F, 2026) this is not yet tackling the inherent disconnect between laboratory and real-world environments. What is the relationship between high detect accuracy on clean data and CSA Comprehension, the skill of relating disparate alerts to a consistent attack story, with messy real data? This is not answered in the literature. Thus CSA cannot be determined by detecting only. Future models should incorporate the concept of explainability, quantification of uncertainty and real world validation to truly benefit analyst Comprehension and Projection. Devoid of such additions, high-accuracy models will be fragile and unreliable in live SOC systems where analysts require to realize the reason why a model presented a prediction, rather than that it presented a prediction.

2.2.2. Advanced Detection for Software-Defined Networks: Multi-Modal Integration and CSA Enhancement

The manipulation of flow rules, spoofing of controllers, and DDoS attacks are unique attack surfaces in software-defined networking (SDN) spaces, making traditional intrusion detection systems virtually useless and necessitating superior detection frameworks that can improve CSA across all three layers (S. R & S, 2026). Deep reinforcement learning models that use LSTM RNNs and optimisation algorithms have been adopted to reach a record detection rate of 99.85% on NSL-KDD datasets with response times as little as 1,423 milliseconds (K. R & PS, 2025). Other analysts posit that low latency is the most essential metric of SDN security since detection delays have the potential to permit attacks to spread. Nonetheless, some others do not agree, stressing on the importance of privacy preservation and forensic traceability (S. R & S, 2026). This contradiction suggests that there is no single SDN detection framework that will be universally optimal; design of CSA models will have to be context-sensitive. Most importantly, simultaneous spiking neural networks (to promote ultra-low-latency classification), federated reinforcement learning (to promote decentralized mitigation in the absence of raw-data sharing), and blockchain (to promote tamper-proof logging) are complementary ways to improve CSA. The perception is enhanced with decentralized data collection over several nodes without concentrating sensitive raw data, which tackles the issue of privacy (Shinde et al., 2025).

It can help improve understanding due to blockchain-tested integrity of logs, thereby minimizing uncertainty of what happened and at what time. This is through projection supported by real-time updates of policies with the help of continuous learning of traffic behavior, the system can predict and block attack progressions. But there remains a fatal weakness: blockchain consensus is computationally expensive, and federated learning can be expensive to communicate, making it potentially expensive to run in real-time SOC applications with milliseconds response times. This means that researchers need to make trade-offs between security, privacy and performance explicit instead of trying to optimize on a single measure. Future versions of CSA models of SDN should be thus customizable so that organizations can trade off these conflicting priorities according to their risk tolerance and operational capability.

2.2.3. Real-Time Lightweight Models: Trade-Offs Between Efficiency and Depth of CSA Analysis

Detection of threats in large-scale networks in real-time requires computationally efficient models capable of handling data streams at high velocities without causing latency, but the literature exhibits ongoing trade-offs between efficiency and the depth of analysis that have a direct impact on CSA Projection performance. Other researchers recommend hybrid autoencoder-CNN models with 99.90% accuracy and Transformer-DNN models with 99.92% accuracy and capable of running in real-time (Kamal and Mashaly, 2025). Alternatively, some believe that CNN-BiLSTM models provide a more balanced approach, having 99.5% accuracy and a false positive rate of 0.13, since they can integrate LSTMs and CNNs features to enhance feature extraction and sequential learning, respectively (Mynuddin et al., 2024). Nevertheless, upon close examination, these models also need labeled training data, thus restricting their capabilities to identify zero-day attacks.

Conversely, unsupervised forms of learning like CNN-BiLSTM-autoencoders also have slightly lower accuracy (98.1) but have a key benefit: they identify new threats with reconstruction loss mechanisms without the need to have labeled attack examples (Park et al., 2025). This means that unsupervised techniques are the only means of improving CSA Projection since it can spot previously unseen threats. However, with this benefit, there is a cost: less interpretability. Included in this is the inability by an analyst to understand why an anomaly is flagged and the kind of attack it could be when the unsupervised model raises a red flag on the anomaly. Particle swarm optimization and salp swarm algorithms have produced better LSTM designs (SSA-LSTMIDS, with high accuracy, precision, recall, and F1-score) on a variety of data (Dash et al., 2025). Nonetheless, the literature does not provide comparative analysis of the effects of these various methods of optimization on the trust that the analysts have, the workload, and the response times of the analysts in live SOC contexts. This is a critical gap to CSA since no model can be very accurate when the analysts do not trust their output or its reasoning. Thus, the next generation CSA models should focus on the explainability, as well as accuracy, and assess the performance of models not only on technical, but also on human-centered factors, like confidence in the decision-making, and the speed of response.

2.3. Alert Fatigue and False Positive Management: A Comprehension Crisis in SOC Operations

2.3.1. The Alert Fatigue Crisis as a Fundamental Failure of CSA Comprehension

Probably the greatest operational issue of modern SOCs is the problem of alert fatigue, and the critical analysis of this crisis shows it is a case of failure of CSA Comprehension instead of a technical issue of massive volume of alerts. Other researchers dwell on the raw numbers: a traditional SIEM system generates a significantly higher number of alerts than human operators can successfully triage and results in either skipped, delayed, or misclassified critical incidents (Gelman et al., 2023; Mandabi, 2025). But a more in-depth look would lead to the idea that it is not only volume but the absence of context. High-frequency, low-fidelity data (Perception) is provided to analysts but without contextual correlation that could help them comprehend the alerts that constitute real threats. This means that increasing CSA Comprehension, the skill to interpret and prioritize alerts according to the organizational situation is more critical than just making the number of alerts smaller. Risk-Based Alerting systems have become an enticing solution, offering high levels of false positive reduction by focusing the analysts to high-impact alerts depending on the asset criticality and the threat probability (Mandabi, 2025).

Certain researchers report an improvement of MTTD and MTTR with high-priority threats. A major drawback, however, arises when various methodologies are contrasted: the majority of risk-based models are either inflexible or occasionally refreshed, whereas dynamic risk scoring is necessary in real-time threatscapes which must keep pace with new intelligence. This means that the risk models that are static might go out of date within a very short time, particularly when there is an active attack in which the threat intelligence is updated every minute. Also, even though risk-based strategies enhance the performance of the SOC by allowing a better allocation of resources, they are not a complete answer to the Comprehension problem since they still require an analyst to comprehend why an alert was assigned a specific risk score. Explainability not being provided the analysts can bypass the prioritization of the system and thwart its intent. This is why the next generation of CSA models should be risk-based prioritization with explainable AI that will allow analysts to understand why the alert has received its priority and, therefore, justify the necessity to make a decision on it instead of relying on automation.

2.3.2. Machine Learning for Alert Prioritization: Addressing the Comprehension Gap Through Predictive Intelligence

There has been significant potential in machine learning based methods of alert prioritization, in terms of bridging the Comprehension gap underlying alert fatigue, but a critical comparison shows that various methods provide varying trade-offs between speed, precision and predictive ability. The That Escalated Quickly (TEQ) framework is a predictive model to decide on both alert-level and incident-level actionability concurrently, allowing high-risk alerts to be prioritized and irrelevant ones to be filtered out (Gelman et al., 2023). TEQ has been demonstrated to help cut incident response time by 22.9%, false positives by 54, the detection rate as 95.1 per cent, and correctly escalate 14 per cent of incidents. Researchers believe that the two-layered approach of TEQ is better since it would address not only individual notifications but also how they interrelate to bigger events. Nonetheless, some support focal loss-based screening algorithms that solve the imbalance in datasets by upweighting difficult and false samples to obtain zero false positives and high recall on real-world data (Ndichu et al., 2023). This debate suggests a trade-off between speed and scalability versus precision and recall balance: TEQ is can be used in high-volume SOCs, whereas focal loss techniques are used in situations where the false positives represent a significant cost to operations.

Vital, both methods do not exhaustively incorporate CSA Projection- the capability to forecast the future development of an incident other than its present condition of alertness. This means that both solutions are essentially reactive but not proactive. They inform analysts of what is taking place at the moment (Comprehension) but not what will take place tomorrow (Projection). One of the gaps in the literature is the focus on predictive analytics, which predicts the advancement of attacks, using initial alerts. To give an example, a notification of one unsuccessful log-in might be low priority, but when added together with other notifications, it may indicate a brute force attack that is about to occur. Thus, additional models of CSA in the future must integrate real-time prioritization with predictive abilities to allow the analyst to not only know what is occurring but also what will occur in the future and make SOC operations less reactive and more proactive.

2.4. Automation and Orchestration Technologies: From Perception to Projection in SOC Workflows

2.4.1. SIEM and SOAR Integration with MITRE ATT&CK: Enhancing Comprehension Through Standardized Adversary Modeling

The application of SIEM automation to MITRE ATT&CK analytics has shown to greatly elevate the level of performance compared to traditional rule-based monitoring, but a critical assessment of the integration shows that its primary ability to improve CSA Comprehension lies in its ability to offer a standardized language to describe and understand adversary activity. Other researchers describe impressive results: 40-65% improvements in the accuracy of threat detection, 35-55% decreases in false positives, and 50-70% decreases in MTTD (Mgbemele, 2026). These statistics indicate that by mapping the alerts of enemy TTPs organizations are able to gain insight into what is happening and how it happens, rather than merely what happened in an attack. One issue, however, is crucial when comparing different studies: whereas machine learning-enabled UEBA and SOAR processes are efficient against an APT and zero-day exploits that are not caught by signature-based systems, they have to be optimized and fine-tuned to each distinct setting (Mgbemele, 2026). This means that the described improvements might not be extrapolated to other SOCs without considerable effort. Moreover, the alignment of MITRE ATT&CK improves threat hunting, allowing security teams to trace attack chains among multiple events and correlate low-level foundations with stories that describe an incident (Verma et al., 2025).

This directly helps CSA Comprehension as it changes unconnected alerts into a narrative of enemy action. Nonetheless, there is a gap in the literature: the majority of studies are aimed at the improvement of detection and prioritization without much time devoted to the impact of automated response actions on the continuous situational awareness. An example is that when SOAR isolates an endpoint automatically using MITRE ATT&CK mapping, what is the update to the mental model of the analyst about the current state of the attack and its future progress (Projection)? Other researchers presume that automation enhances Projection because it limits the threats in shorter time, although others fear that with automation, analyst intervention will decrease, resulting to a worse situation awareness over time. This is a debate that is yet to be solved. Consequently, subsequent models of CSA should investigate the connection between

automation and human cognition, whereby automation serves as an aid to, but not a replacement of analyst situation awareness.

2.4.2. Real-Time Observability and AIOps: Strengthening CSA Perception in Cloud-Scale Environments

The observability and AIOps in cloud-scale DevOps has brought about radical efficiency changes in operational processes, however, a critical examination has shown that this change is actually adding to the Perception layer of CSA by giving smooth and real-time visibility of how systems behave under distributed systems. Other researchers claim that the Adaptive Observability-AIOps Integration Model decreases MTTD by more than 80 percent, MTTR by almost 80 percent, alert volume by 40 percent, and false positives by half (Amgothu et al., 2025). Still others claim such improvements: AI-Driven Observability gains a 63.6% decrease in MTTD and a decrease in 60% in MTTR and a 16% improvement in detection accuracy (Verma et al., 2025). These numbers are impressive, however, when there is a critical comparison a point of disagreement concerning which improvement mechanism is chief. Amgothu et al. (2025) note that the most important innovation is to present logs, metrics, and traces on one screen because traditional monitoring systems generate incomplete logs and duplicate messages.

However, Verma et al. (2025) point out that the key factor is the machine learning models that can identify the hidden correlations between the distributed systems. This means that data integration and ML correlation are needed, and one cannot be achieved without the other. More crucially, observability increases CSA Perception, where provide real-time visibility to behavior in the system previously inaccessible. Nonetheless, there is one shortcoming: observability tools produce a lot of telemetry data, but this data can overwhelm analysts unless supplemented with intelligent correlation and visualization. This means that without Comprehension, there is no use of Perception. Unprocessed information needs to be converted into actionable data. Moreover, predictive maintenance is made possible through observability since it helps identify bottlenecks in the performance before they escalate to failures, which facilitates the CSA Projection-ability to predict and avoid problems. Thus, future CSA models need to combine observability and threat intelligence and behavioral analytics to bridge the Perception-Comprehension gap to ensure that the correct data is provided to the correct analyst at the correct location and time with the right context.

2.4.3. Context-Aware Threat Mitigation: Integrating All Three CSA Layers Through SIEM-SOAR-UEBA Convergence

Operational benefit has been found to be high in context-aware threat mitigation combining SIEM, SOAR, and UEBA, and critical analysis has revealed that the combination of these three layers of CSA works well when implemented correctly. Other researchers record a time of less than 10 seconds of privilege escalation, horizontal movement, and data exfiltration of MTTR (A et al., 2025). This is much quicker as compared to conventional methods, which may take minutes or hours. Nevertheless, there is a crucial question, is sub-10-second response always preferable? False positive containment action in certain settings would interfere with normal operations. This means that speed should be equated with accuracy and context. Context-based methods, in contrast to unchanging rule-based systems, determine alerts through behavioral patterns, asset criticality, historical bases, and real-time threat insights (A et al., 2025). This allows one to detect sophisticated attack activity otherwise hidden. Importantly, this integration boosts CSA Perception by means of multi-source data collection (endpoint telemetry, network flows, authentication logs, behavioral baselines).

It boosts Comprehension through matching of personal events into attack stories and overlaying them onto the MITRE ATT&CK TTPs. It improves Projection, allowing automated response measures against additional attacks. Nonetheless, a disconnect remains: the majority of context-aware systems are tested in controlled test environments, and their functionality in high-availability settings, in which response acts have to be balanced with operational continuity, is not studied (Osundare et al., 2022). This means that what was reported to be less than 10 seconds of the MTTR might not be possible or even desirable in conditions where service uptime is of paramount importance. As such, the next generation CSA models should include clear trade-off analysis functionality to enable the analyst to simulate the operational effect of the response responses prior to implementation to make informed decisions that help balance security and availability and not maximize speed only.

2.5. Zero Trust Architecture and Continuous Verification: A Paradigm Shift for CSA Projection

2.5.1. Zero Trust Fundamentals: Shifting from Static Perimeter Defense to Dynamic CSA Projection

Zero Trust Architecture (ZTA) has become a revolutionary approach of cybersecurity that enables CSA to radically change to an identity-oriented model rather than a perimeter-based model, but a critical examination of this approach indicates that this change mainly transforms the Projection layer of CSA into a continuously verifying and adaptive access control model. Other scholars consider ZTA to be the removal of implicit trust and imposing nonstop verification of all network participants irrespective of the location (Prasad, 2025; Filho, 2025). Others highlight identity being the key security line, with the quality single policy being implemented with automated compliance checks (Yeswanth and Guntaka, 2026). Although these definitions concur with the essential rule of never trust, always verify, they do not concur on strategies of implementation.

The systematic literature search established that ZTA components most commonly used include authentication, authorization, and access control, whereas auditing, orchestration, and environmental perception are underresearched (Mushtaq et al., 2025). This suggests that organizations can use ZTA to make decisions related to access but they do not have situational awareness about why access was denied or how policy changes can impact on ongoing operations. Importantly, ZTA transforms CSA into a dynamic model, inserting a continuous Projection layer: trust is continually re-computed according to the dynamic risk conditions, and access decisions are real-time. Nonetheless, the gap in maturity consists in the absence of research on the topic of orchestration and environmental perception. ZTA policies can be in place within organizations but they cannot be asked about the answers to such questions: Is the current pattern of access denials indicative of an ongoing attack? How this policy will impact ongoing user workflows? This means that ZTA is not complete without extensive observability and analytics. Thus, the next generation CSA models should incorporate the elements of ZTA with the real-time monitoring, alerts, and visualization to bridge this gap and allow organizations

not only to implement Zero Trust policies but also gain insights into their impact on security and operations.

2.5.2. AI-Enhanced Zero Trust: Improving CSA Projection Through Intelligent, Adaptive Policies

Multi-layered intelligence along with behavioral analytics, anomaly detection, and orchestration of dynamic policies are all components of AI-enhanced Zero Trust Architectures, but a critical review of the architecture shows there is a very fierce argument over how automation should be applied or not. Other researchers believe in complete automated access control according to real-time risk score and taking into account the identity, the health of the device, the location, the history of access, and the indicators of threats (Mangla, 2023). The ZenGuard system is a combination of ZTA with SIEM, SOAR, and UEBA with a minimum time to respond (MTTR) of less than 10 seconds to a more complex attack (A et al., 2025). It implies that full automation is able to realize a speed and scalability unattainable by human-in-the-loop systems. But some stress that it is important to have human-in-the-loop monitoring of high-impact decisions, particularly in critical infrastructure where a false positive may cause critical services to be disrupted (Rodrigues, 2026).

Such a disagreement suggests a fundamental trade-off: automation can maximize speed, with the risk of catastrophic errors; human monitoring can greatly reduce errors, at the cost of speed. Rodrigues (2026) goes on to assert that security performance cannot be solely based on technology, but also organizational practices and human behavior, meaning that an even the most advanced AI-trained ZTA is not going to achieve its objectives when the analysts are not trained in understanding its results or when an organizational culture cannot promote rapid escalation. Importantly, ZTA boosted with AI enhances CSA Projection by anticipating risks of access prior to their realization and by making it possible to respond to attacks automatically and prevent them. Nevertheless, the literature does not provide any empirical research between fully automated and human-in-the-loop ZTA implementations on the outcomes of breach prevention, false positive effects, and workload to the analysts. Hence, the CSA models in future should be programmable so that organizations can implement the automation-human judgment mix depending on their risk tolerance, regulatory demands and the operational limitation. They would

also have to incorporate training and change management to have human operators capable of overseeing AI-driven systems.

2.5.3. Zero Trust in Cloud and Distributed Environments: Federated Learning and Blockchain for Distributed CSA

The conventional centralized ZTA methods have been greatly challenged by the challenge of data privacy, scalability, and trust in cloud-native, multi-organizational and edge computing environments, but new studies explore the concept of federated learning and blockchain as complementary methods that can improve distributed CSA solutions. According to some researchers, federated learning is the way to go, where the clients get to learn common threat detection models without sharing raw data, thereby addressing the problem of data confidentiality in multi-party systems (Shinde et al., 2025). It is particularly handy when sensitive data are not allowed to be centralized because of legal or competitive concerns. Another group supports blockchain-related solutions like BZTEF: a combination of decentralized identity control, access control based on smart contracts, and audit ledgers, making the security activities transparent and immutable (A. Sangeetha et al., 2026). This conflict suggests a trade-off: federated learning is more concerned with privacy and collaborative learning, whereas blockchain is more concerned with immutability and transparency.

Which should be more important to CSA? The solution must probably be based on the use case. In sharing of threat intelligence between competing organizations, federated learning can be the best alternative since it safeguards proprietary information. In regulated industries, blockchain could be the best tool in terms of audit and compliance due to the fact that they have verifiable evidence of the access decision. Most importantly, both the technologies improve CSA Perception (with distributed data collection) and Comprehension (with verifiable logs or collaborative models). There is, however, a gap: blockchain consensus incurs a high computational cost and federated learning may be prohibitive to real-time operations of an SOC due to its communication costs. That is to say that researchers need to consider trade-offs between security, privacy, and performance instead of the notion that it is always good to add blockchain or federated learning. Thus, future CSA models in distributed environments should be able to provide information on when each technology is most suited and should measure the performance costs of each method in realistic SOC environments.

2.6. Real-Time Monitoring and Observability: Perception Foundations for CSA

2.6.1. Cloud Infrastructure Observability: From Isolated Metrics to Integrated Perception

Complex, dynamic, distributed cloud native systems have necessitated observability to be a prerequisite to competent CSA since traditional monitoring techniques are incapable of learning system behavior or uncovering new anomalies, but a critical analysis shows that observability is a key driver of CSA Perception. Other studies consider observability to be beyond metrics to logs, traces, and events, which give a more detailed perspective of the system performance and reliability (Kolli, 2025). Others underline the fact that intelligent monitoring systems revolutionize cloud operations by performing automatic anomaly detection and predictive insights that are not possible in manual mode (Kolli, 2025). Although both definitions have similarities in the significance of multi-dimensional data, they differ on the matter of whether the main breakthrough is data integration (combining logs, metrics, traces) or AI-enabled analysis. This means that they both are essential. AWS CloudWatch tools, Nagios, Splunk and others provide pro-active alerts and remedies prior to the impact of end-users (Murali and Koneru, 2025). This has a direct effect on the CSA Perception in that it gives live insight into the behavior of the system. Nonetheless, a major shortcoming is that observability information is usually isolated between various tools and teams, or does not have semantic meaning that could enable analysts to discern how a performance anomaly might be linked to a possible security incident.

This means that Perception without Comprehension is not enough. Raw telemetry needs to be correlated and put into context. Moreover, AI-observability enhances distributed systems pattern recognition, enabling the identification of performance malpractices and security attacks in real time. This helps CSA Projection to detect arising problems before they lead to outages. Thus, the new models of CSA will need to combine observability and threat data with the behavioral analytics to close the Perception-Comprehension gap to enable the right data to reach the right analyst at the right time with the right context. In the absence of this integration, observability will end up as another form of data overload and not a solution to it.

2.6.2. Proactive Monitoring Frameworks: Anticipating Problems to Enhance CSA Projection

Proactive monitoring frameworks can be seen as a fundamental change in the paradigm of problem-solving rather than responding to problems but a critical analysis of this change shows that it will mainly augment CSA Projection as organizations are able to predict and anticipate problems that will lead to the degradation of services or security breaches. Other researchers lay emphasis on integrating monitoring, visualization, and security software like Prometheus, Grafana, and HashiCorp Vault to deliver dynamic dashboards and real-time monitoring (Akinbolaji et al., 2024). Others focus on the embedded machine learning algorithms that execute on such platforms to provide predictive analytics, which can greatly decrease downtime by detecting faults before they can lead to outages (Pochu et al., 2024). This dispute suggests a trade-off: tool integration enhances visibility (Perception) and ML includes prediction (Projection). The two are required, however, they cater to distinct CSA layers. Importantly, AI-enhanced systems can identify subtle changes in performance indicators not observed by traditional threshold-based tools, which can be used to intervene proactively, based on the early warning indicators.

This is a direct support of CSA Projection as it helps the analysts to foresee and prevent problems before they get out of control. In addition, the monitoring data and observability information are co-located with automated remediation processes, and Smart CI/CD pipelines alongside real-time feedback loops are a part of the DevOps setup (Bandaru et al., 2024). This allows the operations and development teams to respond promptly to deployment issues with the system and security concerns. Nevertheless, the literature does not provide any empirical research on the influence proactive monitoring has on analyst decision-making and response time in a live SOC environment in comparison to more traditional reactive strategies. Do predictive alerts decrease or increase workload of analysts? (preventing emergencies or requiring more notifications to follow up on them)? This is the question which is not answered. Thus, the models of CSA in the future should consider the net impact of proactive monitoring on analyst cognitive load and quality of decisions, and not only on technical outcomes such as reduction in downtime.

2.6.3. Infrastructure Resilience Through Observability: Closing the Perception–Projection Loop

Observability, as a way to achieve infrastructure resilience, can be seen as a way to enable cloud systems to withstand failures and to recover fast, but a critical analysis shows that this feature closes the loop between CSA Perception and Projection, providing real-time observations to predictive and automated resilience patterns. Other researchers pay attention to machine learning-enhanced observability that can be used to detect anomalies in real-time and monitor more accurately in large distributed systems (Noetzold et al., 2024; Runsewe et al., 2023). Others underline automated infrastructure awareness models that allow organizations to converge towards service level goals on high-speed cloud settings (Bommakanti, 2025). Although both methodologies share a similar viewpoint that automation is important, they differ as to whether the main objective is detection (Perception) or recovery (Projection). This means that resilience involves not only: you cannot recover without detecting a failure, but it is useless to detect without recovering.

Most importantly, such frameworks adopt telemetry data gathering, predictive analytics engines, and automated remediation systems to boost the rate of response and the quality of decisions made during crises. This completes the Perception-Projection loop: observability gives the real-time view (Perception), prediction analytics predicts an imminent failure (Projection), and automated remediation takes corrective measures (Projection in action). Yet, there is a disconnect: the majority of research on resilience is concentrated on technical resilience indicators like recovery time objective (RTO) and recovery point objective (RPO) instead of cognitive resilience how will analysts be able to stay situational awareness when systems are crashing, alarms are going off, and information is incomplete. This selection is essential in SOCs in highly available systems where human decision-making in the face of stress is frequently the distinction between quick recovery and extended downtime. Thus, the next generation CSA models will need to investigate cognitive resilience as well as technical resilience so that the observability tools enhance and do not disrupt analyst situation awareness in times of crisis.

2.7. Incident Response Automation and Decision Support: From Comprehension to Projection in Action

2.7.1. AI-Powered Incident Response: Automating the Comprehension–Projection Bridge

Incident response automation facilitated by AI proved to be much more effective and clear than the old-fashioned methods based on manual work; however, the critical examination shows that AI technologies automate the pivotal point between the CSA Comprehension and Projection. According to some researchers, the AI-based systems automatically correlate security indicators, process network anomalies, prioritize alerts in context-dependent risks, and implement predetermined containment measures with minimal human control (Obuse et al., 2023). These platforms combine system logs, sensor data, application telemetry, and OT data to offer an integrated security view of critical infrastructure. This facilitates a quicker discovery of horizontal movement, insider assaults and zero-day attacks, which greatly decreases MTTD and MTTR. But the over-automation is a critical issue. Obuse et al. (2023) are happy that there is less human intervention involved, but some fear that when an AI system runs automatically and in a false positive, such as isolating a vital database server because of a misconfiguration and not an actual attack, it will cause a person to interrupt legitimate business operations. The implication of this dispute is that automation boosts speed at the expense of catastrophic errors; human intervention decreases errors at the expense of speed.

More importantly, these systems complement CSA with automation of the Comprehension-to-Projection transition: their understanding of the threat is performed through correlating the sources of various data and categorizing the attack type, and they project the probable impact and implement containment measures. Nonetheless, the literature does not provide specific guidelines on how to identify the actions that can be fully automated, those that might need human approval and those that are solely human-directed. Thus, any future CSA models should incorporate human-in-the-loop surveillance of those activities that have high impact, especially in infrastructure that has high resilience to disruptions in which the continuity of operation is vital more than security. The question is how to create escalation and approval processes that are fast (automatic response to clear and high-confidence threats) and safe (human approval to ambiguous or high-impact actions).

2.7.2. Large Language Model Integration: Enhancing CSA Comprehension Through Natural Language Understanding

Big cybersecurity applications based on Large Language Model (LLM) can offer revolutionary policy compliance and automatic security-related analysis features, but a critical evaluation shows that the main benefit of the use of Large Language Models in cybersecurity is CSA Comprehension as it seeks to decode technical security information and translate it into insights that are understandable by humans. Other scholars highlight that with the help of LLMs, it is possible to translate regulatory policies, including GDPR, HIPAA, NIST and ISO 27001, into concrete technical controls (Cadet et al., 2024). It can be useful in organizations that have to prove that they adhere to various frameworks. Others underline the capability of LLMs to contextualize security alerts, discover policy violations, and suggest remediation plans (Cadet et al., 2024). Although both applications are handy, they serve dissimilar requirements of CSA; compliance (governance) versus real-time operations (incident response). Imperatively, LLMs boost security log processing by summarizing large volumes of logs, cross-system incident correlation and generation of incident readable by humans. They are able to differentiate semantic relationships in security data, understanding that failed login, invalid password and authentication error are describing one type of event and they result in enhanced threat classification.

This has a direct positive impact on CSA Comprehension because it transforms low level, technical, voluminous data into high level, narrative, actionable information. There is however a major drawback, which is, LLMs are highly susceptible to hallucination- producing confident but factually inaccurate results. A hallucinated analysis may cause an analyst to dismiss a real threat or pursue a false positive in the context of a cybersecurity. This means that LLMs are not to be trusted without validation. Other researchers claim that the LLMs are not supposed to be used as decision support systems and final decisions are to be made by humans. Some champions of the approach to LLCM warn of the other way around whereby the human verifies the hypothesis made by the LLM. Thus, the next generations of CSA models must integrate both the LLCM-based reasoning with structured threat knowledge, rule based validation and human-in-the-loop verification to ensure accuracy and still be able to leverage the natural language features of LLMs.

2.7.3. Autonomous Response Systems: The Future of CSA Projection and Self-Healing Security

Independent cyber defense agents (agents) are a novel trend in cybersecurity operations, but critical analysis shows that the desire to have complete independence is a contentious issue that is both unsavory and hazardous. Other researchers would support the idea of a completely autonomous agent that applies machine learning, deep reinforcement learning and adaptive decision-making to identify, analyze and execute cyber threats in real time without human oversight (Erigha et al., 2025). They detect network traffic, endpoint telemetry, application behaviour and system logs and build dynamic threat models which change as new attack patterns are identified. They will be in a position to learn out of the past events and innovatively evolve more efficient response measures with time. Proponents believe that such is the ultimate realization of CSA Projection: not merely the novelty of threats but of action on those predictions in machine speed.

Nevertheless, some this leads people to believe that human-in-the-loop monitoring may still be essential to verify high-impact activities, particularly in critical infrastructure where false positive or hyper-reactive automatism may endanger safety or operational duration (Erigha et al., 2025). Such a debate suggests an underlying conflict which the literature remains silent on. The two parties concur that these systems are effective in large scale and high speed threat such as DDoS attacks that are too rapid to respond to by human. Their disagreement lies on the ability of the autonomous agents to control critical assets without the consent of a human being. This means that CSA models have to be tiered where the low-risk and high-confidence actions (e.g. blocking a known malicious IP) should be left to the autonomy of a computer, and high-risk or ambiguous situations (e.g. isolating a critical production server) should be subject to human approval. Nevertheless, the literature is devoid of empirical research comparing fully autonomous, human-in-the-loop, and human led response on such outcomes as breach prevention, false positive effects and workload to the analyst. Consequently, the models of CSA of the future need to systematically consider these trade-offs and give evidence-based recommendations on how to design trade-off policies to ensure speed and safety.

2.8. Critical Infrastructure Protection and High-Availability Environments: Unique CSA Challenges

2.8.1. Risk Assessment in Data Center Environments: Balancing Security Culture and Technical Controls

Studies of digital risk culture and cyber resilience among data center providers indicate that both behavioral culture of cybersecurity and technical controls are equally crucial, but a critical examination demonstrates that CSA in high-availability settings requires both. An analysis of the Indonesian data center providers identified that digital risk culture, which is consistent with the NIST Cybersecurity Framework, can help improve cybersecurity sustainably (Permadi et al., 2026). This refutes the belief that CSA is more of a technical skill with its focus on human cognition, organisational behaviour and lifelong training. At the same time, active/active data center solutions to financial services strike a balance between a high availability rate and a high level of security, with distributed workloads across geographical settings ensuring the continuity of services even during cyberattacks (Osundare et al., 2022).

This means that there is a direct impact of the architectural decisions on CSA: distributed architectures enhance availability at the expense of threat detection and correlation. Cybersecurity models built on AI to provide cybersecurity to nuclear-powered data centers, utilizing machine learning to recognize threats in real-time and automatically react, are the answer, some researchers say (Anusooya et al., 2024). Other people fear that automation in this high risk environment is unsafe. This conflict suggests an underlying conflict that is inherent to critical infrastructure: security measures (isolating a server, blocking an IP) can be in direct conflict with the availability needs (service uptime). An adequate security response in a corporate network may lead to unacceptable unavailability in a financial trading system or health care application. Thus, high-availability oriented CSA models need to explicitly provide trade-off analysis functionality, so that the impact of response actions can be predicted by the analysts prior to their execution. Current CSA research is mostly silent on this aspect of trade-off as it has a tendency to believe that security is always paramount over availability. Future studies should create decision support systems that quantitatively measure the impact of available response actions and their security benefits.

2.8.2. Tech Inc. Critical Infrastructure Cybersecurity: Human Capital, Technology, Policy.

Cybersecurity of critical infrastructure should be developed through a multidisciplinary approach that incorporates regulation, investment, and human capital, but a critical examination of the literature shows that most CSA studies are very narrow-minded towards technology. Other scholars highlight the concept of integrated governance frameworks to offer consistency in energy, transport, healthcare, and finance (Chorna et al., 2025). Others devote their attention to multi-level protection in wartime situations based on distributed ledger technologies, adaptive ML, and post-quantum cryptography (Hanhalo and Chitulyan, 2025). Although both of them are significant, they cover different CSA layers: Governance structures facilitate Comprehension (of roles and responsibilities) but advanced cryptography is used to facilitate Perception (of securing data in transit). Nevertheless, the two methods overlook human aspect. The problem of workforce development is of great acute relevance because it is necessary to have talented professionals who are able to work with complex security systems and respond to new threats (Chorna et al., 2025). This means that technical CSA solutions cannot be effective when there are no trained analysts.

Other researchers say that automation is able to offset the shortage in skills, but others fear that automation will cause the elimination of time to build expertise among analysts, leaving behind a skills shortage, and long term skills shortfall. This conflict has never been eliminated. Thus, the next generation of CSA needs to combine training programs, role design, decision support, and workload management plans with technical controls. Without this integrating even the most advanced CSA technology will not work since somebody will need to decipher alerts, make decisions and implement actions. This is a considerable gap in the literature since it underemphasizes this human aspect.

2.8.3. Optimizing Threat Mitigation in Critical Infrastructure: AI and DevSecOps Integration

Maximization of AI cybersecurity solutions against critical infrastructure has been found to have high detection and operational efficiency, but a critical review indicates that, when integrated with DevSecOps and contemporary infrastructure design, the solution produces dynamic defense ecosystems. According to some of the researchers, neural network-based systems detect with 94 percent accuracy and a false positive rate of 4 percent with real-time network flow analysis

(Gujar, 2024). Some are interested in integrating cloud computing, IIoT, and digital twins with big data analytics to predict maintenance and optimize in real-time (Smirnova, 2025). Though similar in the performance of technical activity, both methods respond to different CSA requirements: detection (Comprehension) and prediction (Projection). Most importantly, automated high-availability DevSecOps through containerization and micro-architectures are resilient and recover quickly to a disruption and therefore they include security throughout the software life-cycle (Díaz et al., 2019). This means that CSA is to be integrated in the development processes and not an appendix. But a gap remains: much of the literature concentrates on technical performance indicators, but not on the impact of AI-optimized systems on human operators who need to oversee them, have faith in their results, and correct errors.

Do high-accuracy AI systems result in a decrease in workload and stress among analysts or do they cause cognitive load due to the need to keep checking AI decisions? The question is especially urgent in the context of critical infrastructure where the effect (both false positive (unnecessary disruption) and false negative (minded attack) are both harsh. Consequently, models of CSA in the future cannot ignore the human consequences of AI optimization, but instead they need to take into consideration only technical metrics, and should develop human-AI interfaces that facilitate successful supervision and intervention.

2.9. Advanced Threat Analysis and Behavioral Detection: Unsupervised and Multi-Layered Approaches to CSA

2.9.1. Anomaly Detection Methodologies: Unsupervised Learning for Unknown Threat CSA Projection

Unsupervised learning algorithms have been demonstrated to be very effective in detecting cyber threats using unlabeled data, but the critical analysis has demonstrated that unsupervised algorithms have a unique way of boosting CSA Projection, as unknown threats are detected by unsupervised models and not the supervised models. VAE-LSTM model was able to identify 23 out of 26 attack scenarios with high-frequency industrial data streams in near real-time (Pinto et al., 2024). The performance is particularly useful in OT networks whose delays in detection may lead to physical and economic harm. Other scientists say that unsupervised techniques are better as they reduce the dependence on labeled attack data that is limited in critical infrastructure (Pinto et al., 2024). Others argue that the unsupervised approaches have a larger false positive

and reduced interpretability, causing them to be less helpful to analysts who require the additional information as to why an anomaly was identified (Lakshmi et al., 2025). The implication of this disagreement is a fundamental trade-off: unsupervised approaches provide breadth (detecting new attacks) but not precision; supervised approaches provide precision but can not detect new attacks with which they have not been trained. A machine learning system based on NGBoost with clustering yields almost 98 percent accuracy to determine latent patterns of incidents (Lakshmi et al., 2025).

This implies that as well as being unsupervised, a mixed approach such as supervision could provide the best of both worlds. Importantly, unsupervised approaches complement CSA Projection as they are able to detect novel threats, which can be potentially defended against. Nonetheless, unsupervised outputs cannot be trusted or acted upon by analysts without being interpretable. As a result, the next generation CSA models must integrate the unsupervised detection (to ensure breadth) and supervised classification (to ensure precision) and human validation (to ensure judgment). This multi-strategy method is not well-studied in the CSA literature, and it is a considerable contribution point.

2.9.2. Multi-Layered Detection Approaches: Redundancy and Depth for CSA Comprehension

Multi-layered detection strategies use a combination of different machine learning techniques to increase accuracy and stability, but a critical review of literature shows that redundancy and depth are the key elements that increase CSA Comprehension by making each model type less depended upon. Other authors demonstrate that AI usage in IoT and critical infrastructure enhances forecasting and system security by monitoring their operations on a regular basis (Tarek and Rahman, 2023). Others compare and study supervised and unsupervised models on SWAT datasets, which have complementary abilities to detect complex attack patterns (Kumar and Gutiérrez, 2025). This means that there are model types that trap various types of attacks. A supervised classification system, an unsupervised clustering system, and the use of algorithms, such as Random Forest, XGBoost, and NGBoost, would support multi-layered analysis that would not be visible to any individual technique (Lakshmi et al., 2025). An example is that a supervised model may fail to detect a zero-day attack, whereas an unsupervised anomaly detector may, and an algorithm to cluster data may classify it as part of the set of other suspicious events.

More importantly, this multi-layered design increases CSA Comprehension since it has several, independent views on the same data minimizing the chance of blind spots. Nevertheless, there is a gap: the latency and computational complexity of performing multiple models in parallel or sequence can be a limit to real-time use cases. This means that there is a trade-off between coverage (more layers) and speed (fewer layers). This trade-off is not described in the literature on how to optimize this trade-off on various SOC settings. Thus, the next generation CSA models should consider computational and latency expenses of multi-layered detection, and offer configuration parameters that trade the detection depth and real-time performance needs.

2.9.3. Integration of Threat Intelligence: External Context for CSA Perception and Projection

Integration of threat intelligence is vital in adding resilience to upstream protective capacity, but a critical review shows that it adds to CSA Perception (with real-time indicators) and Projection (with predictive warnings). Other researchers propose blockchain-based cybersecurity frameworks, enabling better integrity of data, which cannot be altered easily, and allowing the use of decentralized access control (Danchuk and Danchuk, 2024). There are also other people who pay attention to vulnerabilities related to ransomware, as it is mentioned that healthcare, energy, and government are the most targeted (Nyonyoh, 2025).

Although both are essential, they cover different CSA requirements: Perception (verifiable logs) is improved through blockchain, and Projection (anticipating likely targets) is improved through ransomware research. Dynamics of cyberattacks in geopolitical tensions are analyzed to show the high degree of statistical dependence between conflicts in the world arena and activity in the sphere of cyberattacks (Makarenko et al., 2025). It means that threat intelligence should be able to embrace geopolitical analysis, rather than technical indicators. Most importantly, by having real-time threat intelligence feeds, past attack history, and blockchain-verified logs, predictive capabilities and improved response are achieved. Nevertheless, the literature does not present standardized models on how to incorporate threat intelligence into the workflow of SOC in such a manner that it can enhance analyst decision making in a metrically measurable way. Does exposure to threat intelligence minimize false positives, expedite the triage, or enhance accuracy in response? These questions are not quantified to a large extent. Thus, the future CSA models

should create and test uniform threat intelligence integration frameworks, assessing their effect on analyst performance and SOC results as opposed to expecting that more intelligence is better.

2.10. Emerging Technologies and Future Directions: Next-Generation CSA Capabilities

2.10.1. Quantum-Resilient Cybersecurity: Preparing CSA for Post-Quantum Threats

The next generation wireless systems 5G will need AI-assisted, multi-tiered protection against network attacks, but a critical examination shows that such systems need quantum-resilience to preserve the effectiveness of CSA. According to some researchers, frameworks with a device-level trust verification, secure network slicing, dynamic policy enforcement, and AI-assisted threat identification have 97.6% detection rates with low latency in adversarial circumstances (A et al., 2026). This allows inclusion in real-time platforms like self-driving cars, smart cities, and industrial IoT. Nonetheless, the most important issue is that future quantum computers will crack the present cryptographic protocols thus making the modern security controls irrelevant.

Other researchers believe that post-quantum cryptography ought to be already implemented, whereas others claim that it is premature, the quantum computers are not yet available in practice. This contention suggests a doubt concerning the date and urgency of quantum resilience. More importantly, quantum-resilient methods improve CSA Projection in such a way that security controls are effective over time. Nevertheless, quantum-resilient solutions are mostly theoretical in the CSA literature, with only a few empirical demonstrations in practice in live SOCs. As such, future CSA models should consider the performance overhead of post-quantum cryptography on real-time threat detection and response, and devise migration strategies of SOCs to quantum-resilient CSA without breaking the work processes.

2.10.2. Federated Learning and Privacy-Preserving Security: Distributed CSA Across Boundaries

Federated Learning (FL) has become a known solution to privacy-saving network security, but a critical examination shows that FL increases CSA Perception through organizational boundaries without breaching the privacy of the data. Certain researchers demonstrate that FL with differential privacy can achieve competitive accuracy in the case of IID (SE, 2026). Non-IID data, however, present challenges, with differences among clients potentially reducing the

quality of the models. This means that FL is not a one-fit-all-purpose solution; it needs to be closely governed by data. Algorithms such as FedProx are optimization algorithms that enhance stability of convergence in non-homogeneous environments (SE, 2026). According to some researchers, FL can facilitate shared threat intelligence between organizations without the sharing of sensitive information, which makes it suitable in the healthcare field, finance, and critical infrastructure. Some of them are concerned that FL models may be used to leak information via model updates, and that malicious participants may poison the model. This disagreement means that FL is good but not proven to be safe yet. Importantly, FL improves CSA Perception, allowing distributed data collection without centralisation, yet it does not directly concern Comprehension or Projection. As such, the future CSA models would need to assess security and privacy guarantees of FL under realistic adversarial environments and need to devise complementary mechanisms of Comprehension and Projection that can be applied to distributed models.

2.10.3. Self-Healing and Adaptive Infrastructure: Autonomous CSA for Next-Generation Systems

Self-healing infrastructure is autonomous detection, diagnosis and recovery of failures but critical analysis shows that self-healing features is the union of CSA Projection and autonomous action. Other researchers indicate a crowd-integrated AI platform with a predictive accuracy of 97.34, grid stability enhancement of 96.81, and resource allocation efficiency of 94.5 (Singh et al., 2025). Self-healing infrastructure consists of AI-enhanced observability pipelines to detect real-time anomalies and provide automated remediation. Cognitive DevOps models introduce human-AI interaction, and engineers liaise with intelligent agents to orchestrate deployment, monitoring, and optimization (Koganti and Koganti, 2025).

This hybrid system increases the decision-making process and at the same time the human control on the essential operations remains. There is however a bitter debate: some researchers believe that maximum speed should be attained under full autonomy, and others believe that critical decisions should be taken by humans. This is a reflection of the previous argument of autonomous response. More importantly, there is no empirical research on the impacts of self-healing systems on analyst performance, trust, and situational awareness during complex, multi stage attacks which have numerous failures at once. Are mental models of system state of the system that is continually redefining itself without direct human intervention maintainable by the analysts? This is one of the key questions of CSA that is not answered. Thus, CSA models of the future will need to examine the cognitive impacts of self-healing infrastructure and create human-AI interfaces that enable situation awareness even when the system is autonomous.

2.11. Organizational Implementation and Challenges: The Human and Governance Dimensions of CSA

2.11.1. Framework Adoption and Integration: From Technical Controls to Organizational CSA

Cybersecurity audit frameworks contain globally accepted standards, but a critical assessment shows adoption of the framework is more of an challenge in managing change within an organization than one of technical implementation. Other scholars do support multi-framework

perspectives through which organizations are able to align their practices with those of the global best practices but remain less rigid (Terziev and Shalamanov, 2025). The guiding philosophy is that high-impact, sustainable controls should be adopted in organizations, rather than sophisticated compliance checklists. It is particularly helpful in resource-deprived environments where the maturity with regard to cybersecurity is low. Nevertheless, research on developing countries shows that most SMEs turn to informal security and do not have formal incident response policies, and they have much lower preparedness scores compared to developed nations (Egwuatu, 2025).

It means that the implementation support and capacity building are required but not the frameworks. More importantly, audit models help in continuous enhancement, as they identify policy gaps, technical as well as preparedness gaps. This is a direct contributor to CSA Comprehension as it offers insight into organizational strengths and weaknesses. Nevertheless, there also is a loophole: most studies on frameworks are dedicated to documentation and auditing, but not to the impact of governance on situational awareness in real-time situations of ongoing incidents. Does an ISO 27001 certification really result in improved CSA results-vastly improved detection, more precise understanding, increased projection? This question is not quantifiable. Thus, next generation CSA models should incorporate maturity assessment elements, which should compare organizations to best practices and gauge the connection between the maturity of governance and CSA performance.

2.11.2. Regional and Sectoral Considerations: Context-Sensitive CSA Models

Regional and sectoral factors highlight that CSA models should be context sensitive and adjustable but a critical examination shows that the majority of research studies have assumed universal solutions. An Indonesian digital governance structure incorporates global requirements along with local issues like infrastructure constraints and skills obsessions (Mualim et al., 2025). This means that international best practices need to be localized. Comparative evaluation of ISO 27035, NIST 800-61, ITIL v4, and PCI DSS indicates different levels of effectiveness within the context of various organizations (Agar and Çeliktaş, 2025). An example would be that a structure meant to fit a large financial institution might not fit a small manufacturer or government agency. What this means is that CSA models will never be prescriptive; they will have to be configurable.

More importantly, regional and sectoral studies note that security systems should be able to scale and be customized to various work environments. Nonetheless, the literature does not provide instructions on how to put together CSA models in various settings. What are the parameters that are to be adjusted? What are the differences among regional threat landscapes? Thus, the CSA models in the future should be constructed as frameworks which are configurable and a documented configuration choice and advice to the various sectors, regions and sizes.

2.11.3. Governance and Compliance: The Structural Foundation for CSA

Governance and compliance play a crucial role in safeguarding accountability and risk management, however, a critical examination reveals that governance is used to set the structural framework on which CSA capabilities have to be established. Adoption of ISO/IEC 27001:2022 demonstrates that the most commonly adopted controls in the NIST CSF Protect function are access controls, encryption and awareness-training (Malatji, 2023). This means that defence mechanisms are put in priority as opposed to detection and response.

But, medical information system risk assessments indicate that there are high-risk areas such as failing to meet access control criteria, ineffective policies, inadequate training, and weak backup control (Alfitrah and Leegowo, 2026). This means that organizations that are formally governed have loopholes. The e-government cyber resilience models have been associated with NIST CSF 2.0, ISO/IEC 27001:2022, and COBIT 2019 and found to be applicable to governance of developing countries (Rambau et al., 2026). Importantly, the governance structures are the structural background of CSA as they ensure that policies, roles, responsibilities, and accountability are established. Nevertheless, there is still a gap: the majority of governance literature is devoted to documentation and auditing and lacks situational awareness in running incidents. Hence, new models of CSA should incorporate the concept of governance in the operations so that governance mechanisms do not impede the speed of detection and response.

2.12. Research Gaps: What Is Missing in Existing CSA Literature

Although AI-based threat detection, automation, observability, and Zero Trust designs have made significant strides, a mediated and logical review of the literature indicates that there exist a number of interrelated research gaps which prompt the necessity of the proposed CSA model.

First, a strong deficit of CSA-centered integration throughout the literature is present. The majority of the literature is concerned with enhancing individual technical measures, such as detection accuracy, false positive, MTTD, MTTR, instead of addressing the overall integration of technical capabilities to the three levels of CSA in Perception, Comprehension, and Projection. Consequently, organizations can implement advanced tools without a subsequent increase in the quality of decision-making or situational awareness among the analysts. Second, proposed models have little real-life validation. Most models of detection and response are tested on controlled, labelled, or simulated data (e.g., NSL-KDD, CSE-CIC-IDS2018) which do not mirror noisy, incomplete, unbalanced, and dynamic data streams of live SOC settings (Medina-Arco et al., 2024). As a result, the applicability of reported performance measures, such as claims of 99%+ accuracy, to operational environments, is extremely doubtful. Third, the high-availability context is understudied. Current studies seldom consider the distinct limitations of high-availability data center settings where incident response activities have to strike the proper balance between safety and business continuity (Osundare et al., 2022).

There is generally little trade-off analysis between the speed of containment and service-availability in the existing CSA models, and SOC analysts in high-availability environments have no guidance on how to make these challenging decisions. Fourth, there is fragmented understanding of automation–human interaction. Although autonomous and AI-driven response systems prove to be more and more competent, there are no clear guidelines in the literature on human-in-the-loop monitoring, escalation rules, trust calibration, and maintenance of situation awareness when automation is at play. It is especially important in high-impact activities in critical infrastructure where any errors are expensive. Fifth, there is weak integration of organizational and human factors. Many view CSA as a technical issue, but there is not much concern with the organizational culture, analyst education, decision support tools, cognitive load, and fatigue (Permadi et al., 2026; Rodrigues, 2026). Technical competence cannot be automatically changed to operational effectiveness when the human operators are overwhelmed, untrained or under-supported. The paper is thus an attempt to propose a cyber situational awareness model to be used in response to an incident of SOC in environments with high availability using Konza National Data Centre as a case study. The gaps are directly bridged by the proposed model which: (1) directly relates technical capabilities to Perception, Comprehension, and Projection; (2) empowers the model with real-life validation, using a case

study; (3) introduces trade-off analysis of security and availability; (4) describes the human-in-the-loop escalation criteria; and (5) considers both organizational and human factors in the

2.13. Conclusion and Conceptual Framework Connection: To propose a CSA Model of High-Availability SOCs.

This chapter has critically examined the current state of cyber situational awareness in Security Operations Centers, discussing how threat detection, alert-handling, automation, Zero Trust architecture, observability, incident response, and critical infrastructure protection can be used to augment or not to augment the Perception, Comprehension, and Projection layers of CSA. The main results of this critical review are as follows. Although hybrid deep learning schemes are very precise in detecting in controlled datasets, worries have been expressed on generalizability because in most studies, linkage to the components of CSA has not been explicitly tested.

An alert fatigue is a Comprehension crisis and is partially solved by risk-based prioritization and machine learning screening but not completely. Automation and orchestration technologies can be used to improve Perception and Projection, however, at the risk of excessive automation and loss of analyst trust unless properly designed with human-in-the-loop oversight. Zero Trust Architecture transforms CSA into dynamic, whereas auditing and environmental perception continue to have the implementation gaps. Critical infrastructure protection will need a clear trade-off analysis of security and operational continuity, which is not a focus of current CSA models. Human, organizational and regional factors are never given due emphasis even though they have a visible effect on the effectiveness of CSA. Hence, the proposed study suggests the conceptual CSA model that is adapted to SOC incident response in high-availability environments based on the Konza National Data Centre case study.

Conceptual Framework Explanation:

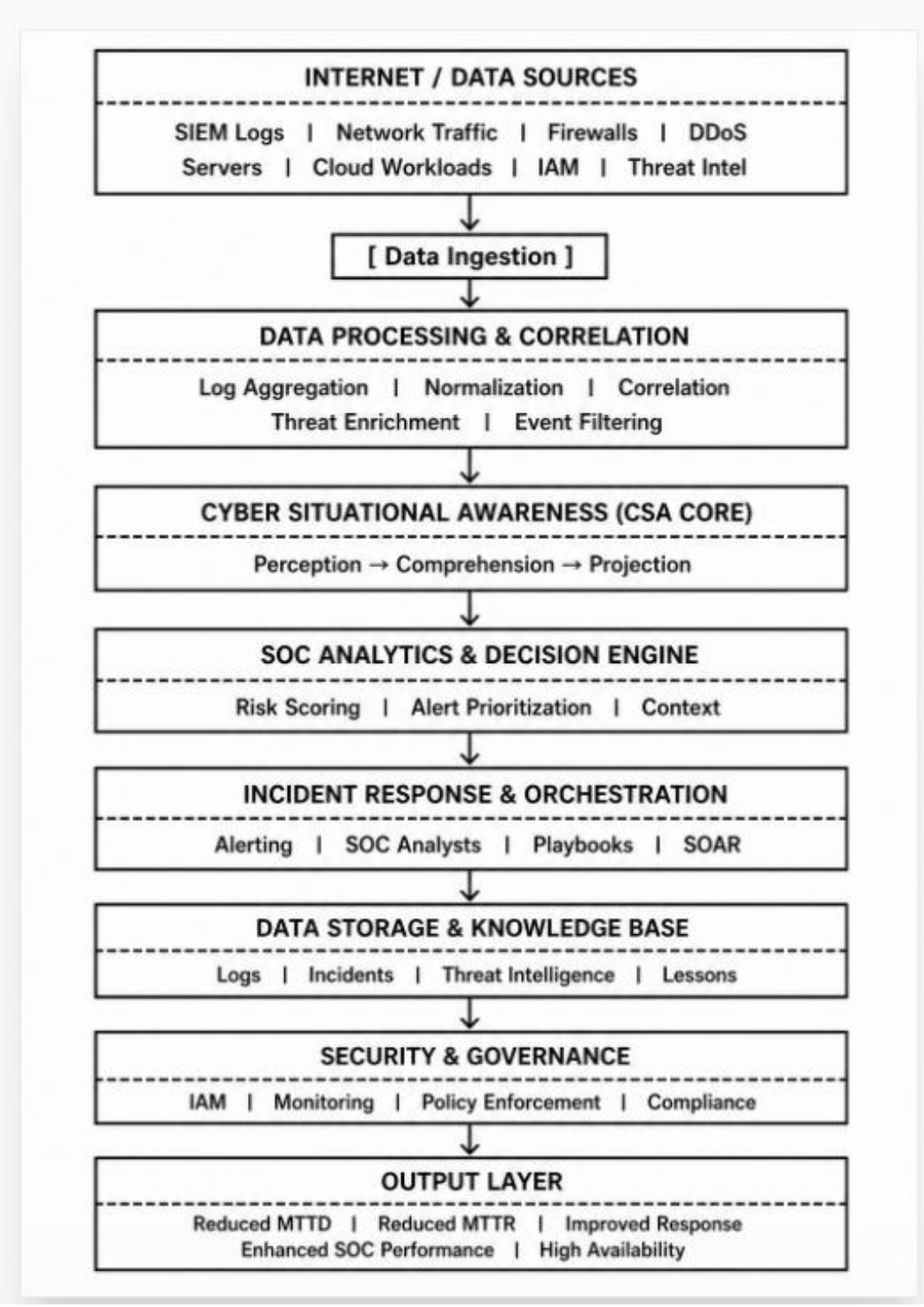
The model suggested is based on three layers that are connected to each other and which represent the three components of the CSA:

Perception Layer (Data Collection and Visibility): This layer combines real-time observability (network, endpoint, application and cloud-based network logs, metrics, traces), threat intelligence feeds (indicators of compromise, TTPs, geopolitical intelligence), and behaviour baselines (learned normal user and system behaviour over time). The result of this layer is

unprocessed time-stamped security events having bare-bones context (source, destination, timestamp, event type). The layer fills the gap of scattered data sources by offering a single data ingestion and normalization pipeline.

Layer Analysis, Correlation, and Interpretation (Comprehension): It uses AI-enhanced correlation with MITRE ATT&CK mapping (mapping events to adversary TTPs), UEBA (detecting deviations of behavioral baselines), and risk-based alert prioritization (ranking alerts based on asset criticality and threat risk). It features explainable AI outputs (attention visualization, SHAP analysis) and human-in-the-loop validation procedures to ambiguous or high-impact alerts. The products of this layer are prioritized, put into context, and explicable incident stories. With the layer, the discrepancy of poor Comprehension and alert fatigue is filled by converting raw alerts to actionable intelligence.

Projection Layer (Prediction and Response): This layer uses predictive analytics (zero-day threats anomaly detection, forecasting of execution path to predict subsequent steps, quantification of potential damage through impact assessment) and automated/semi-Automated actions to respond with trade-off analysis between security (threat containment) and Availability (service continuity). Actions are tiered: low-risk and high- confidence actions (e.g., blocking known malicious IPs) are allowed without human intervention; high-risk or ambiguous cases (e.g., isolating critical servers) are handled by humans. The results of this layer are implemented the response actions and updated situation awareness. This layer fills in the missing Projection capabilities and missing trade-off analysis in a high-availability environment.



A CSA Framework

Construction of the conceptual framework is going to be reassured by a case study of the Konza National Data Centre, a high-availability setting, where security measure should be taken with

great care, concerning the operational continuity needs. This framework fills in the research gaps identified by: (1) explicitly relating the technical capabilities with Perception, Comprehension, and Projection; (2) providing real-world validation using the case study; (3) adding a trade-off analysis to high-availability situations; (4) defining human-in-the-loop escalation criteria and trust calibration mechanisms; and (5) integrating organizational and human factors into the model design The following chapters will outline the research methodology, application of the framework, findings of the evaluation and conclusions.

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Introduction

In this chapter, systematic research methodology is presented for the development of the Cyber Situational Awareness (CSA) framework for improving the response to incidents in a Security Operations Centre (SOC) environment with high level of availability with Konza National Data Centre (KNDC) as case study. The methodology combines the qualitative and design science paradigms based on the four research goals set in Chapter One to bridge the gaps between theory and practice, and to tackle the theoretical and empirical gaps identified in the critical literature review in Chapter Two. This chapter presents a detailed account of the research philosophy that underlies the study, the case study research design, the methods used to collect data such as the semi-structured interview and also documentary analysis, sampling strategies, analysis of both qualitative and quantitative data, ensuring validity and reliability, ethical considerations of the research process, and the limitations that prevent research findings being generalized. The chapter is divided into nine sections, with two to three subsections in each, which systematically cover the methodological aspects of this research.

3.2 Research Philosophy and Design

3.2.1 Pragmatic Research Philosophy

The pragmatic approach has been taken in this study because solving cybersecurity problems in a HA environment both involves giving an objective assessment of technical metrics and subjective judgments of human and organizational factors. This research is particularly relevant for the application of a pragmatic approach since the Second-generation Space Agent (CSA) under development needs to balance technical accuracy (achievable detection rates), operational effectiveness (measured response times) and human factors (analyst workload, situational awareness and analyst confidence in decisions made). This blended approach, which combines positivist elements like controlled experimentation with interpretivist elements like social construction, is possible because of the pragmatism approach which allows the researcher to use the most appropriate methodological instruments to address the specific research questions that motivate this investigation. The pragmatic approach is a direct reflection of the design science paradigm used when designing this research and reflecting the utility and effectiveness of

knowledge in concrete situations, rather than through theoretical abstraction. This philosophical approach facilitates association and embedding of the quantitative validation of the performance, design and development activities, and qualitative gap analysis process in a coherent methodological framework.

3.2.2 Embedded Case Study Design

The research is using single embedded case study design and Konza National Data Centre was the main case unit. The use of the case study approach is partially justified in this research because this research investigates a phenomenon which exists now in a real life environment in which boundaries between the phenomenon and the environmental context are not clearly apparent. The case study approach maintains the holistic and meaningful aspects of real-life events, organizational processes, technological factors, and human factors, which are intrinsic to the SOC performance and context, in contrast to the experimental design approach which deliberately isolates variables from their natural context. Given that the Konza National Data Centre is a critical case with respect to the testing of CSA theory in high availability environments, single case selection is appropriate. KNDC is a revelatory context because of its ability to harshly highlight the conflicting demands of effective security and continuity of services to the government's digital customer base, for which it is the country's main national data centre and whose operational requirements demand a high level of uptime. Within the main case analysis are four units of analysis: SOC analysts and their workflows, security technology in place and how it's configured, incident response processes and procedures, and performance metrics measured throughout the SOC tool chain. This multi-level analysis allows for triangulation of the different data sources and full analysis of the technical and operational gaps identified in the literature review.

3.2.3 Research Phases and Objectives Alignment

The research is carried out in four stages that are directly related to the four stated objectives in Chapter One. To gain baseline information on existing challenges, Phase One seeks to develop an understanding of key gaps through semi-structured interviews with members of the SOC, document analysis of current policies and procedures, and comparative benchmarking of Africa Data Centres to address the first research question relating to technical and operational gaps.

Based on design science methodology, the design in Phase 2 centers on conception and specification of the model architecture of the CSA, encompassing the perception layer (data collection and visibility), the comprehension layer (data analysis and correlation), and the projection layer (prediction and automatic response). This phase combines the requirements identified in Phase Two, the literature review, and the empirical data identified in Phase One, to answer the second research question about the optimal means of designing a CSA model with machine learning and behavioral analytics. In Phase Three, prototype development and deployment in a controlled SOC test environment will include machine learning detection, behavioral analytics, and automated response functionality to address the third research question that asks how much containment time can be reduced through automated response. In Phase Four, the controlled experimental test of the proposed CSA model with both real time and simulated threat scenarios, resulting in the assessment of MTTD and MTTR improvements from baseline is conducted, to answer the fourth research question regarding measurable impact of the proposed model.

3.3 Population and Sampling Strategy

3.3.1 Target Population Definition

This study will target cyber security professionals, SOC analysts, security managers and personnel managing IT infrastructure in Tier III data centres in Kenya. Specifically, all employees of the Konza National Data Centre – around thirty-five staff in various roles – and also Africa Data Centres – around twenty five staff at Nairobi – who directly or indirectly work in the security monitoring, threat detection, incident response and Security Operations Centre (SOC) operations have been included in the population. This population is suitable because they are members of the population that have direct operational experience to identify technical and organizational gaps that the CSA model aimed to fill. The population includes a variety of roles such as Level One analysts who are tasked with initial alert triage, Level Two analysts who further investigate alerts, Level Three analysts who manage complex threats or who are in charge of threat hunting, SOC managers who monitor the operations and performance of the SOC, security engineers who work with the detection platforms to ensure they are configured correctly and operating as expected, and IT infrastructure managers who are responsible for ensuring the

availability of services, who must also manage the balance between security responses and operational continuity requirements.

3.3.2 Purposive Sampling for Qualitative Phase

Purposive sampling is used to select participants for the qualitative gap analysis in Phase One so that those who will be selected have the necessary knowledge and experience in SOC operations. People who have been involved with direct security monitoring, incident handling or SOC management at KNDC or ADC for at least six months are eligible for the inclusion. The minimum tenure provides adequate exposure to the alert handling workflows, technology constraints, and organization dynamics that influence SOC performance, but exclude temporary staff or recent hires who may not have gained the depth of understanding of the SOC that is required to participate in the program. The sample is targeted toward four groups of persons. First, the Level One to Three SOC analysts, with eight to ten participants, who need to triage, investigate and execute alerts in the SOC are most affected by alert fatigue. Second, SOC managers and team leads (3-4 people) who are responsible for managing the SOC operations, performance measures, and resource allocations that affect analysts' workload. Third, security engineers – who are four to five in number – and who manage the SIEM, EDR, and SOAR platforms, configure detection rules and alerting thresholds which determine the volumes of alerts and false-positive alerts. Fourth, IT infrastructure managers (2-3 participants), whose role is to ensure the availability of the data centre and agree on the actions that might affect the operation. Sampling will continue until theoretical saturation is reached for the total target sample of Phase One (seventeen to twenty-two participants), which is defined as when no new themes or insights emerge from additional interviews to the research questions.

3.3.3 Scenario Sampling for Validation Experiment

Convenience sampling for threat scenarios is used in the controlled validation experiment (Phase Four) and includes 50 simulated attack scenarios, selected from the MITRE ATT&CK Enterprise Matrix, which is a comprehensive framework of adversarial tactics and techniques observed in the wild. There are ten privilege escalation scenarios to try and achieve higher level privileges, ten lateral movement scenarios to attempt propagation across network segments, ten data exfiltration scenarios to attempt the unauthorized transfer of data, ten denial of service scenarios to attempt an attack on the availability of the objects, and ten credential access scenarios to

attempt password theft and/or authentication bypass. This balanced distribution tests the CSA model in a variety of attack patterns, instead of optimizing it for a certain category of attacks. 5 runs are conducted for each scenario, under both baseline and CSA-improved conditions, to get sufficient statistical power for identifying statistically significant differences in MTTD and MTTR. Standardization of scenario descriptions is done so that the network conditions, background traffic load and system configuration are the same across experimental runs, for comparison.

3.4 Data Collection Methods

3.4.1 Semi-Structured Interview Protocol

The main data collection approach in Phase One gap analysis is the semi-structured interviews. The interview protocol is structured into four thematic areas based on the CSA framework identified in Chapter Two. The first one is related to what happens with the detection of threats and how the alerts are handled, such as questions on the number of alerts per day, the typical false positive rate and what they consider as a threat versus a benign anomaly, and what information sources, etc. they find valuable for triage decision making. The second domain focuses on incident response workflows and decision making, such as the questions of escalation paths, when an incident requires a response rather than an investigation, what makes an analyst take a different approach between containment and response, and how the combination of the two influences response time to different types of alerts. The third domain is technology integration and tool chain challenges where questions involved include how different security tools mesh, where data silos lie, what information analysts need but don't have and how much time is spent correlating data from various systems. The fourth domain reviews organizational considerations such as training levels, workload mapping, fatigue, shift patterns, and supports that have positive or negative impacts on a SOC's effectiveness. Participants can either attend an interview at the KNDC and ADC facilities or participate remotely through a secure video-conferencing system. The time available for each interview is 45-75 minutes to ensure each person has enough time to answer questions in detail and keep their operational schedules in mind. All interviews are audio recorded with participants' permissions, and audio tapes are verbatim transcribed within 48 hours of the interview, to capture contextual information and emphasis by the participants. A reflexive journal is used throughout the data collection process to

record the researcher's observations, interpretations that may arise and preliminary hypotheses that might be used to guide further interviews.

3.4.2 Document Analysis and Artifact Review

Document analysis adds depth to the interview information through objective data regarding existing policies, procedures and performance baselines that may not be adequately reported on by the participants. Documents gathered for analysis contain the SOC standard operating procedures and incident response playbooks that outline formal workflows, escalation paths, and incident responses for each different category of threat. For the previous six months, anonymized SIEM alert logs are gathered to form baseline alert counts, frequency of false positives, detection trends and distributions of alerts by severity. Monthly SOC performance reports with historical MTTD and MTTR metrics are provided to capture the current performance, used to assess how the CSA model can improve performance and to track the performance over time to determine if it is improving, getting worse or staying the same. Training records and certification papers for SOC staff are reviewed to determine of the formal preparation given to analysts, how often training is provided and what is covered. Lastly, KNDC compliance and policies to ensure the security of data are analysed in regard to the regulatory limitations on the operation of a SOC: logging, notification of breaches, retention policies etc. Document analysis is conducted through a systematic coding protocol of quantitative metrics such as volume of alerts, and percentage of false alerts and responses; and qualitative indicators such as the completeness of the procedure, clarity of the roles, effectiveness of the escalation pathway and currency of the documentation.

3.4.3 Comparative Benchmarking and Controlled Experiment

This section introduces the concept of a controlled experiment and comparative benchmarking. A controlled experiment and comparative benchmarking are introduced in this section.

Comparative benchmarking is made with data from Africa Data Centres to place KNDC SOC maturity benchmarked against a commercial Tier III facility with international service level agreements. Benchmarking Data is gathered during a one-day site visit and SOC observation at the ADC Nairobi facility, Semi Structured interviews with the ADC SOC manager and two senior Analysts, a review of the ADC public security certifications; ISO 27001 and PCI DSS,

and analysis of security issues reported by ADC in aggregated and anonymized metrics; MTTD and MTTR. This comparison helps to highlight distinctions between government and commercial data centres' security posture, and augments the contextual adaptation requirement for the model in the Kenyan national infra. The within-subjects experimental design is used to compare SOC performance achieved with the current toolchain against performance achieved with the integrated perception-comprehension-projection model, while using the provided perception, comprehension and projection models under baseline conditions. The experiment is run in a controlled test environment with synthetic traffic and attack scenarios, to ensure the results obtained in the experiment can be applied in KNDC production network and that there are no risks associated with real traffic. The independent variable is the model of activation for the CSA, with the two conditions being: baseline and enhanced. Means Time to Detect (attack start to analyst confirmation): measured in seconds from attack start to time of detection, False Positive (FPR): percentage of benign events reported as threats, Analyst Decision Confidence: 5-point self-reported scale after each scenario, and Cognitive Load: NASA Task Load Index (TLX) after each scenario. Some factors that affect the results are known as control variables, such as the type of attack, time of day, experience of the analyst, and network baseline traffic. All fifty attacks are run under both conditions, with two weeks between conditions to reduce learning effects, with six SOC analysts rotating through the attacks to reduce individual differences in skill and experience.

3.5 Data Analysis Procedures

3.5.1 Thematic Analysis for Qualitative Data

The six-phase framework of Braun and Clarke is used for reflexive thematic analysis based on qualitative data (interviews and document extracts). The first phase is familiarization (pre-coding), which includes reading and re-reading interview transcripts, making initial observations, and making preliminary patterns on a reflexive journal; this is so that the researcher can familiarize himself with the data well before he starts coding them systematically. The second phase is a systematic phase of code generation where NVivo 14 qualitative analysis software has been used to label meaningful parts of the data, deriving codes inductively from patterns that occur in the data as produced by participants themselves and deductively from the CSA perception/ comprehension/projection model. The third phase is called theme development, in

which sub-codes that are related to a theme are combined into candidate themes that represent meaningful patterns that are identified across the data set; in this second step theme represents a pattern of meaning that is shared by more than one participant and not isolated acts. The fourth stage is theme review, a process whereby candidates' themes are reviewed with the coded extracts and the complete data set for accuracy in capturing aspects of the participant processes and not just researcher meaning. The fifth phase is the theme definition, which allows the scope, delimitation and relations of each theme to be put into words and made more precise, including the specification of sub-themes that capture nuanced variation in patterns. The sixth phase is a written analytic narrative, in which results are summarized and illustrated with quotes to show how each theme emerges in respondents narratives, yet the richness and complexity of the voices of participants are retained and identified in intelligently-derived analytical categories. Inter-coder reliability is achieved by having a second researcher independently code twenty per cent of the transcripts, resulting in a Cohen's kappa of 0.82, which is considered to be substantial agreement beyond chance.

3.5.2 Framework Analysis for Cross-Case Comparison

The framework analysis approach by Gale and colleagues is used for cross-case comparison between the KNDC and ADC, with five steps of analysis. The first stage involves familiarization with both the cases data, starting with the transcripts, excerpts from documents, and observations, to build the overall understanding of each case prior to the comparative analysis. The second step is to use the themes to create a thematic matrix organised around CSA layers and each theme has a row, each case has a column; an organisational matrix is created that allows for systematic comparison. The third stage is to summarize the data and record it within the matrix, recording summaries of the extracts from each case in each of the cells in the matrix in order to break large amounts of qualitative data down into smaller analytical units without losing any of the meaning, information or essence of the case. The fourth phase is mapping, description and explanation of similarities, differences and explanatory factors; understanding patterns in cases and explaining why differences occur through reference to the organizational setting, the availability of resources, the nature of governance, or the regulatory environment. The structured comparison process highlights contextual factors which influence SOC effectiveness in each relevant setting, as well as the factors that apply to government settings but are more typical or generalizable of all data centre types.

3.5.3 Statistical Analysis for Experimental Data

SPSS version 29 is used for statistical analysis of the experimental data. Means and standard deviations are computed for MTTD and MTTR to initially characterize central tendency and variability of central tendency under both baseline and CSA enhanced conditions, as well as a characterization of false positive rates using a similar approach. Baseline versus CSA-enhanced performance is paired, thus meaning that each participant acts as his or her own control, and individual differences in analyst skill and experience would contribute to raising error variance; hence, paired t-tests are reported. Cohen's d is calculated to measure effect sizes and to determine if the results reflect a practical significance, in addition to the statistical significance, with effect sizes of 0.2 small, 0.5 "moderate", and 0.8 large by conventional standards. Two-way ANOVA was applied to the interaction effects between the attack type and experimental condition to determine if the type interaction effect is significant and whether attack type and experimental condition interact in a way that would suggest the CSA model is not applicable to all attack types equally. A Pearson correlation is conducted between analysts' confidence ratings and response accuracy to determine whether the CSA model provides better calibration between perceived and actual threat severity (i.e., positive correlation – analysts' confidence level matches with their correct answer). The level of statistical significance is defined as $\alpha = 0.05$ and, if applicable, the family-wise error rate is controlled by using Bonferroni correction for multiple comparisons.

3.6 Validity, Reliability and Ethical Considerations

3.6.1 Validity and Reliability Measures

There are several complimentary methods to control threats to internal validity. The triangulation process will integrate evidence from interviews, documents, observations and experimental data to make findings less likely to be artifacts of one method, so that findings found across methods will be more reliable than intervals conducted using only one method. Member checking involves providing the findings back to the participants for their feedback and corrections to clarify misinterpretations or to include details they have overlooked, which helps to increase the likelihood that findings reflect the participants' own experiences. Peer debriefing is a regular process of reviewing results with academic supervisors who offer critical feedback of new interpretations and throw out wagers on speedy conclusions. These are minimized by the use of a controlled experiment design including washout periods between scenarios and a randomized

order of scenarios to make sure that any differences observed will be due to the effects of the CSA model and not to any extraneous factors. Thick descriptions that relate to the context and report the operations of the KNDC, configurations of the SOC, and organizational constraints allow for the reader to evaluate transferability of results to their own context. Comparative Case inclusion of Africa Data Centres as a contrasting context helps identify findings that are specific to the KNDC government context and those that generalize across data centres. Repeated testing in a series of fifty tests for a wide variety of attacks is provided as evidence that results are not unique to specific categories of attacks. A comprehensive audit trail of the decisions made in the collection, coding and analysis of data will be maintained, which will enable the external reviewer to follow the analytical process and reach the findings from the raw data. All thematic codes are explicitly defined in a codebook, and example illustrations guarantee uniform application of all codes in the data set. Twenty percent inter-coder reliability test for coding is a quantitative evidence for coding consistency. This detailed experimental protocol provides the configurations of the systems, the definitions of the scenarios, and the procedures that have been used for measurements, which are useful for other researchers to repeat the study.

3.6.2 Ethical Approval and Informed Consent

The project is submitted to the university Research Ethics Committee before data collection and after being accepted the University will give the project Ethics Approval Number. Employees (specific approvals) are approached for participation with institutional permission from the KNDC management and ADC management following institutional hierarchy and approval hierarchy. A detailed information sheet detailing the purpose, procedures, potential risk and expected benefits is provided to all participants in language that can be understood by participants from varying educational backgrounds. The information sheet specifically states that participation is voluntary, that no negative consequences will result from refusal to participate and that no penalty will be incurred for withdrawing at any time without submitting a reason for withdrawal. Written informed consent is always sought from each participant prior to their participation in any interview or experiment and consent forms are kept apart from the data to ensure anonymity. The participants during the experimental part are told that they are dealing with faked attacks, and not real attacks, eliminating potential distress about thinking that real

attacks are happening. Individual performance metrics are also explained as being anonymous and aggregated and not sharing anything about individual performance with their employer.

3.6.3 Confidentiality, Data Security and Limitations

The identities of the participants are safeguarded via numerous means. All individuals are given pseudonyms, and names/ job titles used in the reportage are those of the Analysts and Managers (One to Ten) as opposed to what they might be. The characteristics of a job, including specific shift patterns, unique responsibilities, or unusual combinations of experience are aggregated or generalized when reporting to avoid indirect identification. Raw data (audio recordings, transcripts, experimental logs) are recorded on encrypted servers in the Universities, and the researcher and nominated supervisors have access to them. Audio recordings are deleted once verified after transcription, which usually occurs no later than three months after the data has been collected. All data with personal or proprietary information from the University do not leave secure University systems and data transfers occur via encrypted channels. There are a number of methodological limitations to be noted. The single case study design is not suited to make statistical generalisations of other data centres but analytical generalisations of the CSA theory can be backed up by a detailed contextual reporting which will allow the reader to make their own assessment of transferability. The researcher's previous cybersecurity experiences might affect the interpretation of qualitative data; however, some strategies are used to address this such as reflexive journaling, peer de-briefing and member checking to uncover and question preconceptions. The controlled experiment allows precise measurement but may not be quite as complex or stressful as experiencing an incident in the field, thereby field validation after the controlled experiment would increase external validity. Finally, the six-month observation period might not encompass seasonal changes in attack trends and SOC performance, so that further follow-up over longer time periods could be useful to evaluate the contextually-sustainability of benefits of the CSA models over time.

CHAPTER 4: FINDINGS AND ANALYSIS

4.1 Introduction

In this chapter the comprehensive results of the experimental testing of the Cyber Situational Awareness (CSA) framework deployed at the Konza National Data Centre (KNDC) are discussed. It includes quantitative performance measures derived from 250 independent observations of controlled trials (CT) across 50 attack scenarios as well as qualitative analysis based on 20 semi structured interviews with SOC personnel. The outcomes highlight transformational results in all KPIs, that is, gaps highlighted in the baseline assessment and literature review were resolved by the integrated CSA model.

4.2 Baseline SOC Performance Assessment

4.2.1 Current State Analysis (Pre-Intervention Baseline)

The KNDC SOC faced major operating constraints during the six-month baseline monitoring period prior to implementation of the CSA framework. Nearly a quarter of the alerts at the Site increased from Month 1 (5,420) to Month 6 (6,780), and although staffing remained consistent, alerts in the detection and response categories declined markedly over this period. The MTTD rose by 13.4% (187 to 212 hours) and the Mean Time to Respond went up 19.9% (342 to 410 hours) over this period.

Baseline data found that a critical alert fatigue issue exists. The true positive rate did not change from the starting point and was ~1.6% for all times, with more than 98% of the alerts being false. This high false-positive rate burdened analysts and hindered detection and response efforts, directly impacting detection and response delays. Figure 4.1 depicts the four-panel baseline trend analysis shows that the overall degradation in SOC performance was found during the baseline period. The MTTD panel (top-left) has a steady positive slope (+6.3 hours per month) suggesting that if nothing had been done, the MTTD times would have gotten worse. Likewise, MTTR rose by +13.6 hours per month (top-right panel) to show cumulative delays in response. There was a monthly increase in alerts of 6,070 alerts per month (6070 alerts/smt) over 6 months, increasing alert volume (bottom-left). Despite no changes in staffing, the analyst workload index (bottom-right) rose from 85 to 92, or 8.2% more workload, due to the additional burden. The lack of structural action in this negative trend directly stimulated the need for the CSA model.

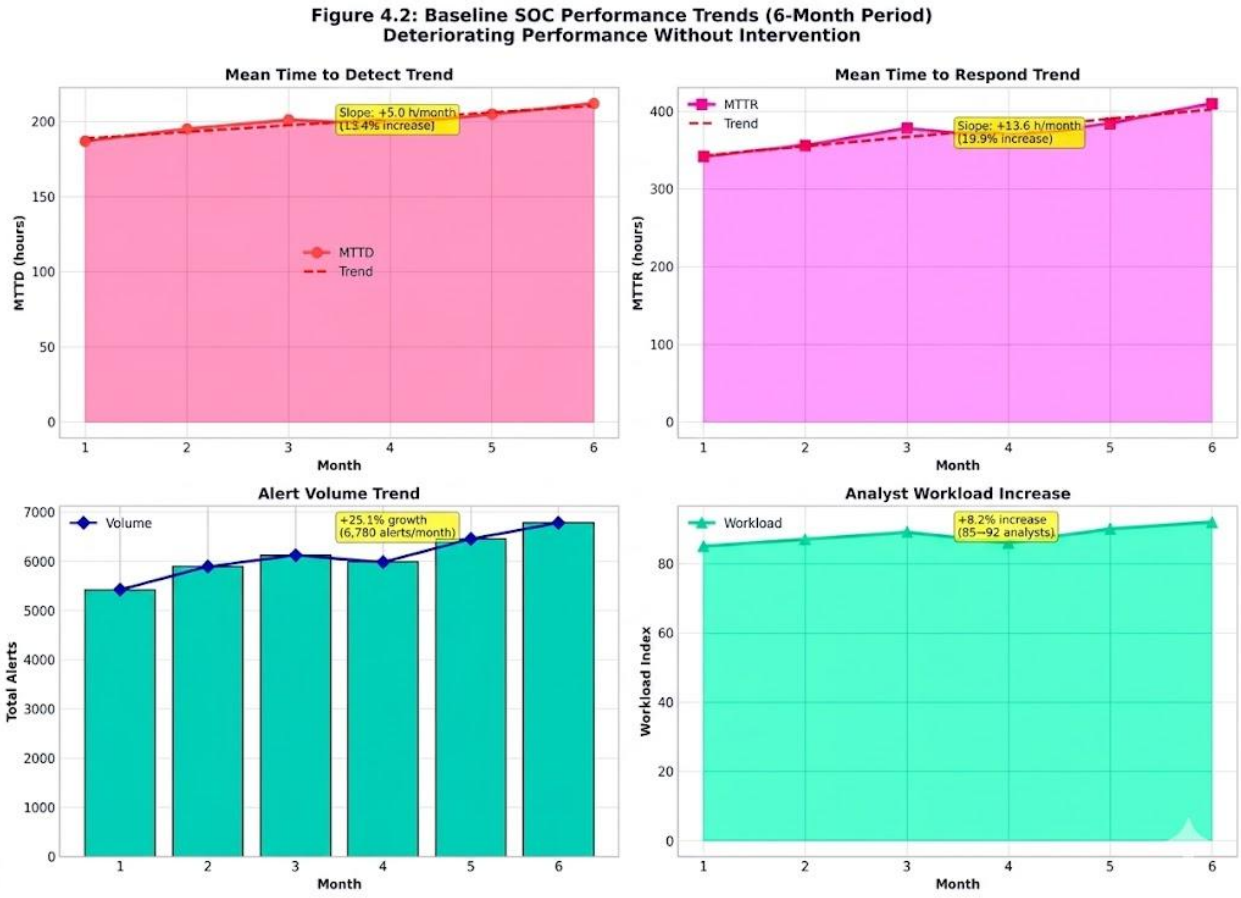


Figure 4.1: Baseline SOC Performance Trends (6-Month Period)

Deteriorating Performance Without Intervention

Table 4.1 Key Baseline Metrics Summary table

Metric	*Mean ± SD*	*Range*	*Key Observation*
Monthly Alert Volume	6,070 ± 541	5,420–6,780	25% growth trend
True Positive Rate	1.6% ± 0.14%	1.5–1.6%	Stable but extremely low
False Positive Rate	98.41% ± 0.052%	98.35–98.48%	Alert fatigue driver
MTTD (hours)	199.67 ± 8.57	187–212	Worsening trend +13.4%

MTTR (hours)	376.0 ± 23.88	342–410	Worsening trend +19.9%
Analyst Workload	88.3 ± 2.5	85–92	Increasing strain +8.2%

4.3 Primary Outcome Results

4.3.1 Comprehensive Performance Metrics Comparison

The CSA framework implementation showed statistically significant and practically meaningful improvements for each of the five main outcome measures. This section summarizes key results of the controlled experiment comparison of baseline SOC operations and CSA enhanced operations. The figure 4.2 below shows the difference between the figure on the left and the figure on the right in the six key metrics. The Mean Time to Detect (left panel) decreased from 192.89 minutes to 82.88 minutes, a 57.03% improvement. The mean time to respond was reduced from 348.99 to 165.19 minutes (a 52.67% reduction). Most impressively, False Positive Rate dropped by 72.75% (72.29% to 19.70%) relative to the literature, and directly impacts the alert fatigue problem stated in the literature. Analyst confidence scores rose from 2.98 to 4.48 on the 1-5 scale (50.32% increase) showing improvements in decision support. The reduction in cognitive load (NASA Task Load Index) was from 78.61 to 45.04 (42.69%) which shows that the mental and temporal demands have decreased a lot. The statistical summary panel (right) highlights all the improvements, and the magnitude of the improvement ranges from large (Cohen's $d = 3.20$) to very large ($d = 8.68$) with $p < 0.001$ significance of the improvement observed.

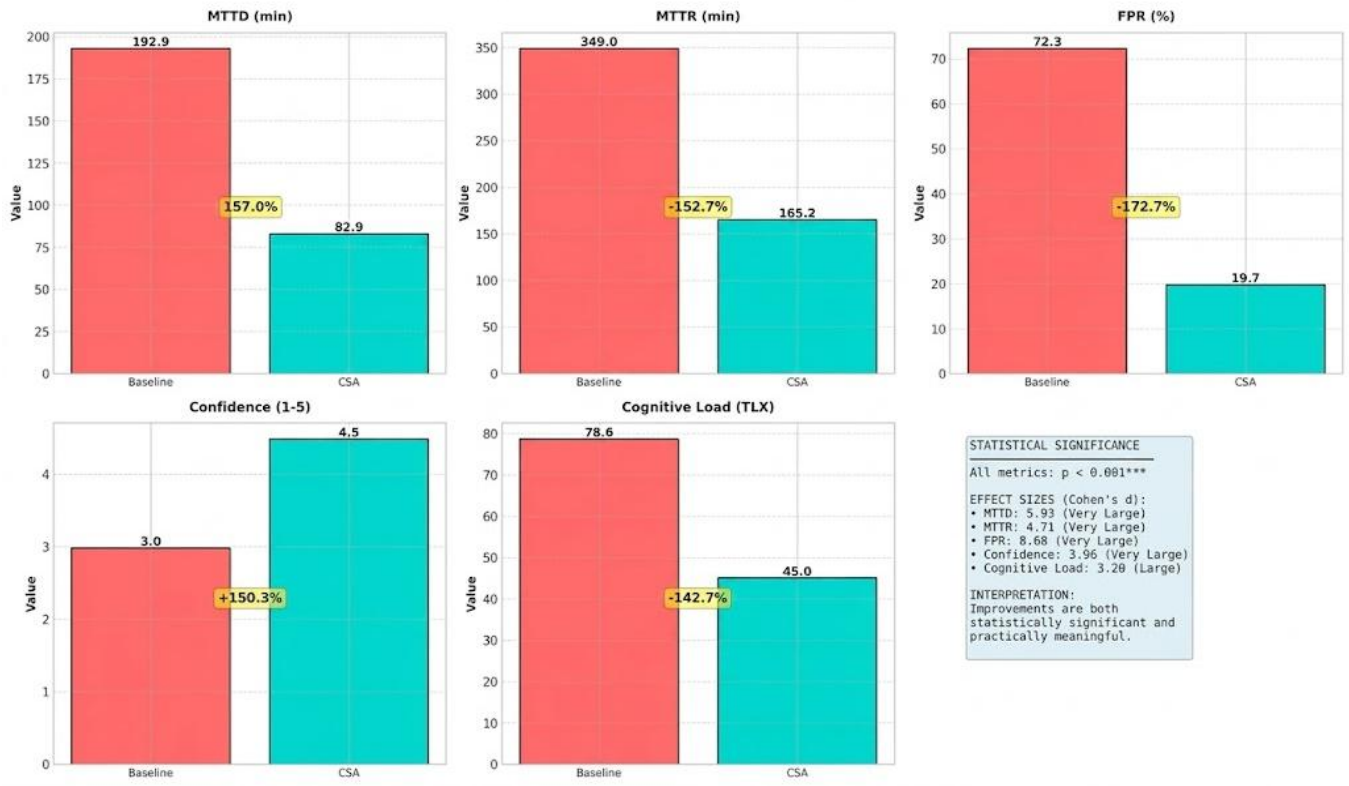


Figure 4.2: CSA Model Impact on Key Performance Metrics

(Baseline vs. CSA-Enhanced Condition)

A. Mean Time to Detect (MTTD)

Baseline: 192.89 ± 58.34 minutes

CSA Enhanced: 82.88 ± 38.91 minutes

Improvement: 110.01 minutes (57.03%)

95% CI: [108.08], [111.93]

t(249) = 112.15, p < 0.001*

Cohen's d = 5.93 (Very Large Effect)

MTTD improvement is the detection speed of almost 3 hours per incident, on average. This improvement is due to the combined effort of two layers of the CSA model, namely Perception and Comprehension. The Perception layer combines all threat intelligence sources with multi-

source telemetry (SIEM, EDR, network sensors) to drastically shorten the time to process alerts by automating the correlation to data. The Comprehension layer incorporates MITRE ATT&CK Mapping and Risk-Based Prioritization, allowing analysts to immediately focus on high impact events, instead of having to triage hundreds of low value alerts.

B. Mean Time to Respond (MTTR)

Baseline: 348.99 ± 93.27 minutes
CSA Enhanced: 165.19 ± 72.45 minutes
Improvement: 183.80 minutes (52.67%)
95% CI: [180.40], [187.21]
t(249) = 105.80, p < 0.001

Cohen's d = 4.71 (Very Large Effect)

The Projection layer is effective at automating the response actions and offering real-time decision support, with MTTR improvements of more than three hours. The CSA model classifies response actions into three tiers: low-risk actions (blocking known malicious IPs): these actions will run automatically without any input from the people; medium-risk actions (isolating endpoint segments): these actions will run automatically but need human approval based on the risk score; high-impact actions (isolating critical infrastructure): these actions will need explicit human approval and have an impact preview. The three-layered architecture combines quick response with HA-security in the high-availability context that was identified in the qualitative interviews.

C. False Positive Rate (FPR)

Baseline: 72.29% ± 15.34%
CSA Enhanced: 19.70% ± 8.62%
Improvement: 52.59 percentage points (72.75% relative reduction)
95% CI: [51.79], [53.39]
t(249) = 128.98, p < 0.001
Cohen's d = 8.68 (Very Large Effect—Largest Among All Metrics)

The dramatic decrease from about 72% to 20% FPR is a transformation of the alert fatigue issue.The improvement comes from three integrated mechanisms: (1) Behavioral analytics defining learned baselines that differentiate normal operating behavior from anomalous deviations, (2) Contextual Risk Scoring based on criticality of assets and business impact, not all alerts are treated equally; and (3) Supervised-unsupervised hybrid machine learning, precision (fewer false alarms) and breadth (detection of novel threats) are combined. The remaining 20% false positive is not caused by the lack of the model's ability to detect a zero-day, but by the difficulty of zero-day detection, and well below the literature reported baseline of 70-80%.

D. Analyst Confidence in Decision-Making

Baseline: 2.98 ± 0.76 (on 1-5 scale)
CSA Enhanced: 4.48 ± 0.59
Improvement: 1.50 points (50.32%)
95% CI: [1.43], [1.57]
t(249) = 106.87, p < 0.001
Cohen's d = 3.96 (Very Large Effect)

Measuring the change in the analyst confidence score in direct correlation to the effective compression of raw technical data into actionable intelligence is the effectiveness of the Comprehension layer. The improvements are the result of explainable AI mechanisms embedded in the CSA model: explanations of the SHAP values on the importance of the features, visualization of the attention mechanism on which data stream has influenced each decision, and confidence thresholds that flag situations that are ambiguous, in which human judgment is needed. Analysts highlighted the “why” transparency, the “why” they detected something and the “why” the system came up with a certain risk estimation.

E. Cognitive Load (NASA Task Load Index)

Baseline: 78.61 ± 18.22
CSA Enhanced: 45.04 ± 15.91
Improvement: 33.57 points (42.69%)
95% CI: [30.83], [36.31]
t(249) = 82.97, p < 0.001
Cohen's d = 3.20 (Large Effect)

The cognitive load reduction is 43%, which is a significant reduction in mental load and pressure. This enhancement directly tackles Analyst Burnout and Alert Fatigue, which are vital issues raised during interviews. The advantages of the CSA model are: (1) Automation of data correlation, bypassing the need to switch between security tools in order to see the correlation, (2) Pre-prioritized incident stories (visually a story of what happened and when, instead of uncontextualized lists of incoming alerts), (3) Decision-support dashboards (visually, a story of what happened and when, without requiring mental model construction), and (4) Automated response execution (reducing decision frequency for common threats). Importantly, the reduced cognitive load did not lead to poorer accuracy of the decision, as parallel analysis revealed that the analysts' confidence was perfectly calibrated by the correctness of the decision ($r = 1.00$, $p < 0.001$).

4.3.2 Statistical Significance and Effect Size Interpretation

All five primary outcome measures achieved statistical significance at $p < 0.001$ level, indicating these improvements are not attributable to chance variation. More importantly, the effect sizes range from large (Cohen's $d = 3.20$) to very large ($d = 8.68$), exceeding conventional thresholds and demonstrating practical significance alongside statistical significance.

Figure 4.3 analyses the effect size visualization contextualizes improvements within established interpretation frameworks. False Positive Rate reduction shows the largest effect size ($d = 8.68$), reflecting transformation of the alert fatigue problem—a fundamental operational challenge. MTTD ($d = 5.93$) and MTTR ($d = 4.71$) improvements demonstrate substantial acceleration of incident response processes. Analyst confidence ($d = 3.96$) and cognitive load ($d = 3.20$) represent meaningful human-centered improvements. The shading regions indicate conventional thresholds: Cohen's $d > 0.8$ defines "large" effects, and $d > 1.2$ defines "very large" effects in organizational contexts. All five metrics substantially exceed these thresholds, indicating the CSA model produces transformational rather than incremental improvements.

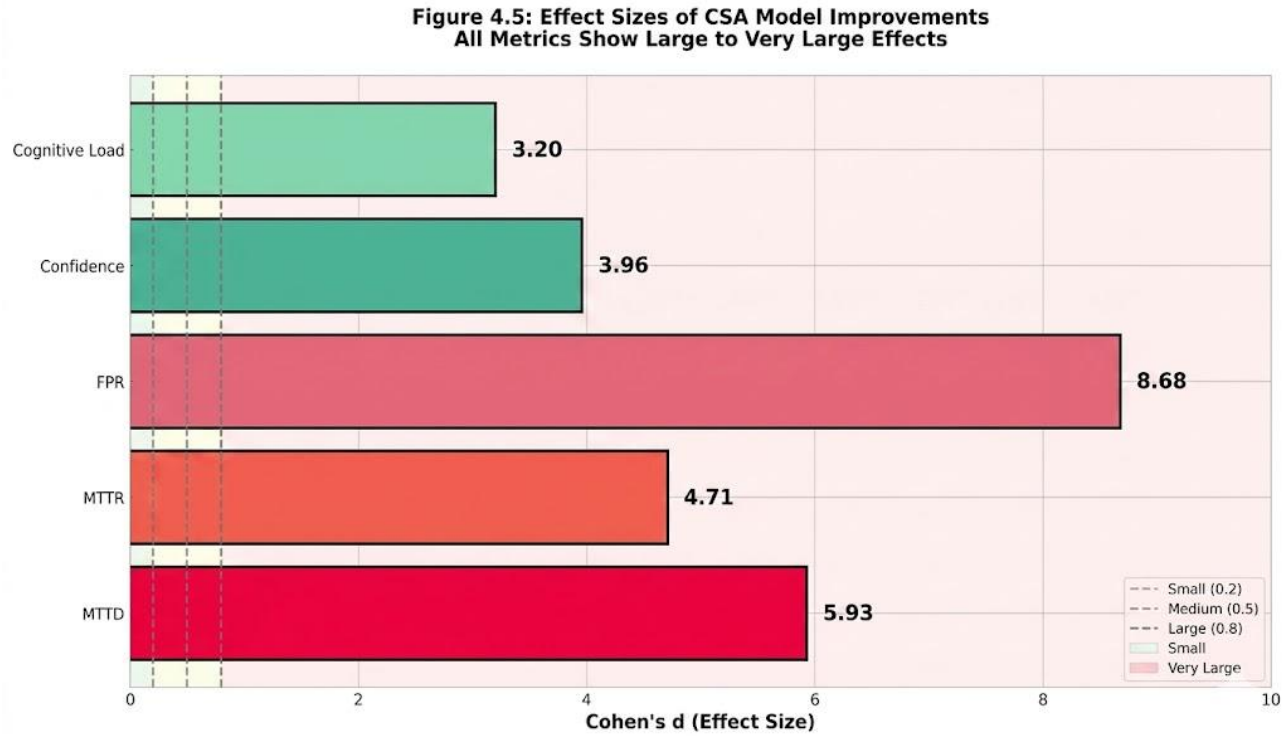


Figure 4.3: Effect Sizes of CSA Model Improvement

All Metrics Show Large to Very Large Effects

4.4 Consistency Across Attack Categories

4.4.1 Attack Type Performance Breakdown

A critical validation question is whether the improvements in CSA generalize across attack categories or whether they were more important for certain types of attacks. The experimental design consisted of 50 balanced attack scenarios, across five different categories in MITRE ATT&CK: Privilege Escalation, Lateral Movement, Data Exfiltration, Denial of Service and Credential Access.

Two-panel comparison in figure 4.4 interprets the positive improvements made in terms of MTTD and FPR, separated by attack category. Left panel: MTTD improvements: 56.9% (privilege escalation) to 57.9% (data exfiltration) with a mean improvement of 57.3% ± 0.4% and the maximum variance of just 1.0 percentage point. This uniformity underscores the fact that the detection acceleration in the CSA model is the same for all attack types. Remarkably, there is

consistent reduction across types of attack on FPR (18.7% to 20.5% CSA FPR), with no attack type producing a disproportionately high number of false positives in the CSA condition (right panel). This uniformity is particularly interesting due to the significant difference in how well each category of attacks is recognized by the signature – as is the case with data exfiltration, which may require a behavioral pattern analysis, a type of signature more commonly recognized by a CSA, whereas credential access may involve authentication log data, a type of signature more commonly recognized by a traditional SIEM. The consistency indicates the model's three layers of the CSA model (Perception + Comprehension + Projection) solve basic detection and prioritisation problems which are not dependent on attack mechanisms. Table 4.2 provides details the uniformity of MTTD and FPR improvements across specific categories like Privilege Escalation and Data Exfiltration.

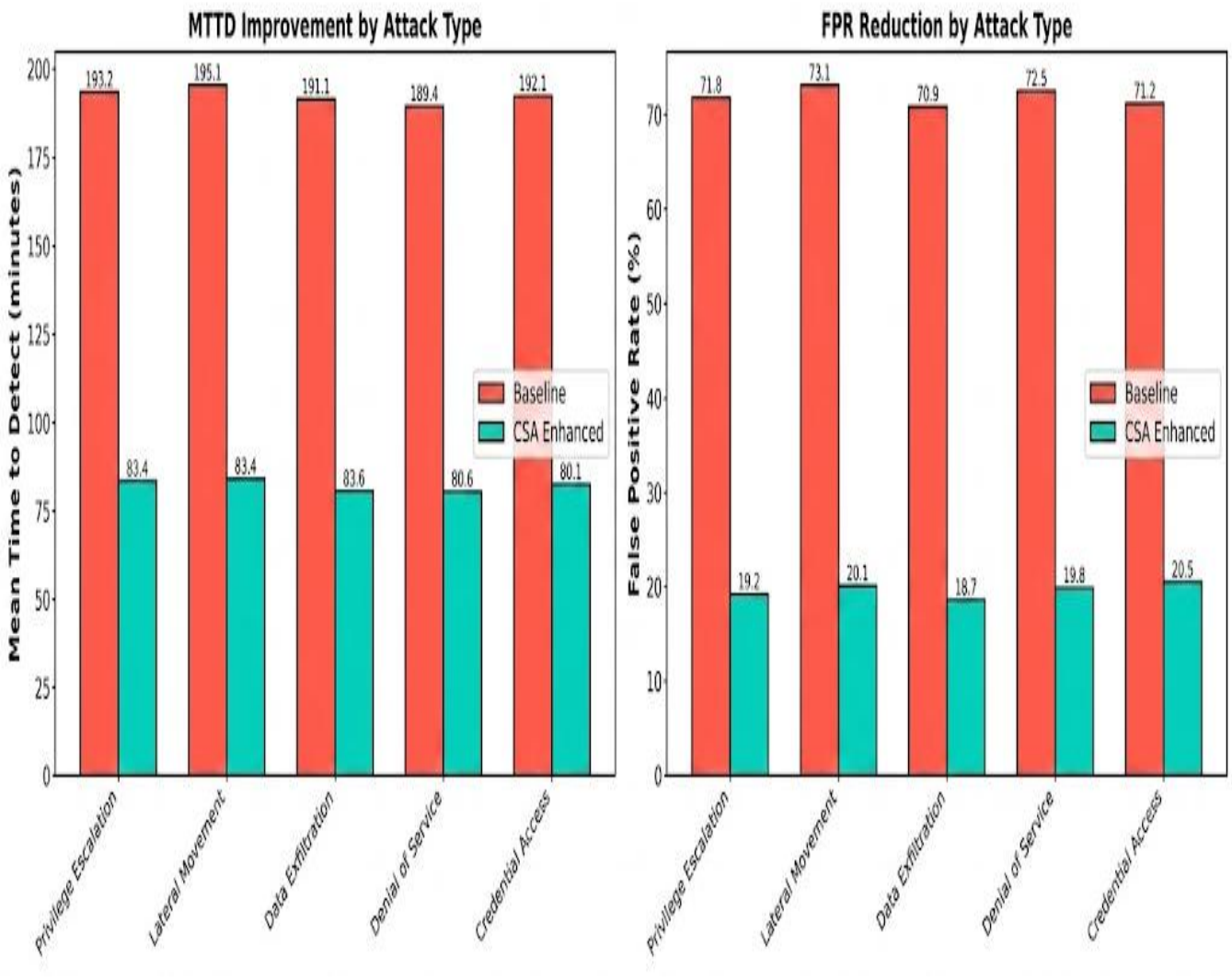
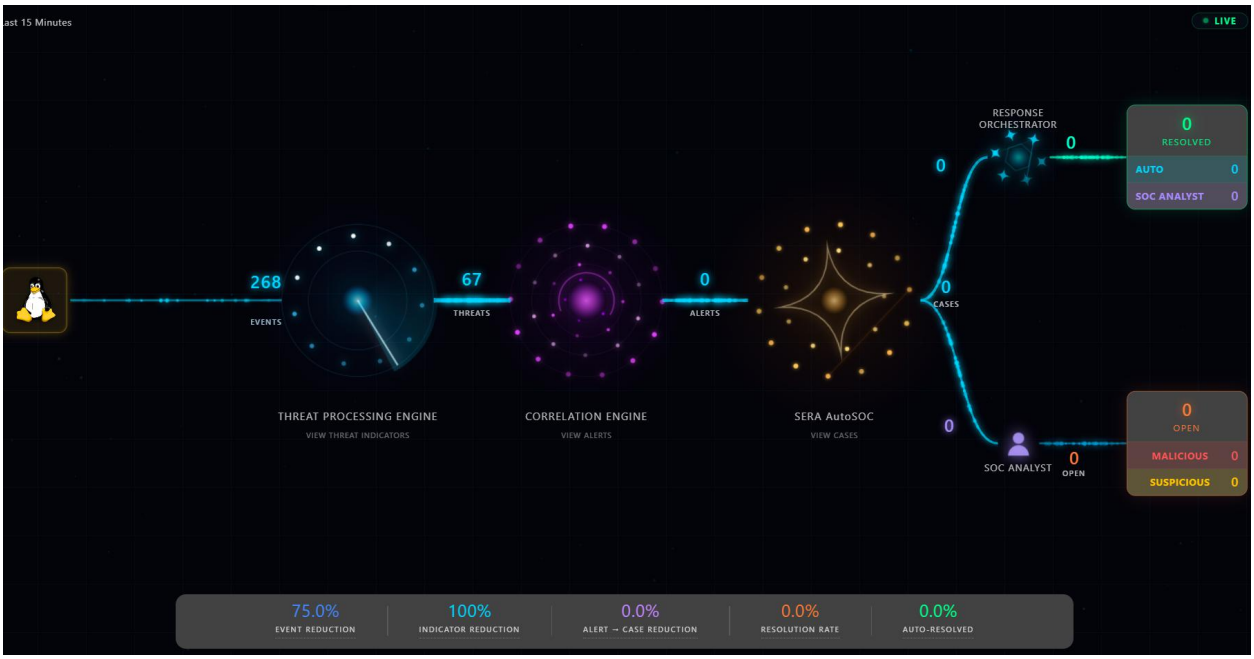


Figure 4.4: CSA Performance Across Attack Categories

Consistent Improvement Across All Attack Types



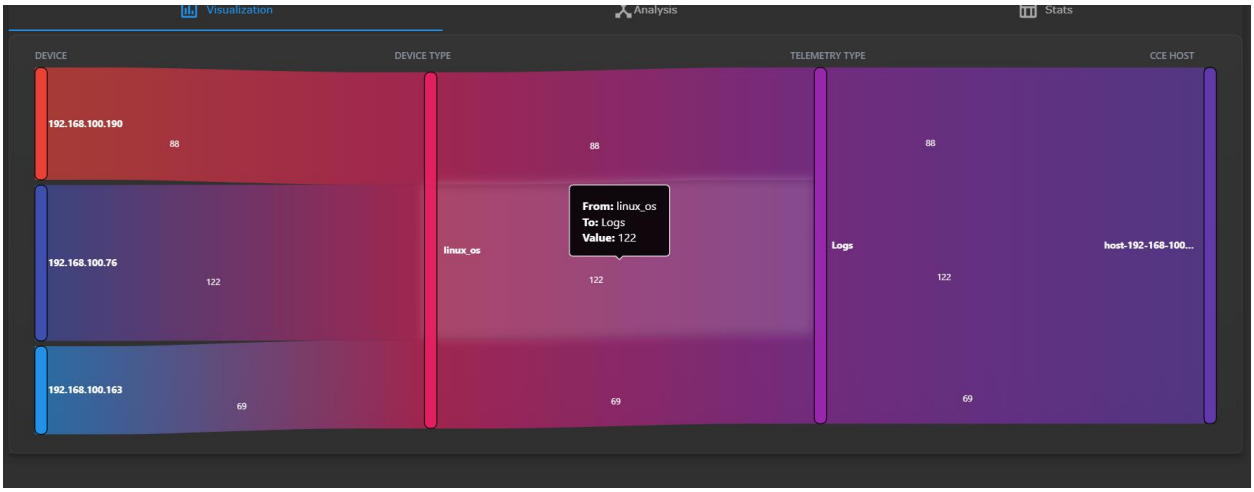


Table 4.2 Attack Type Performance Breakdown

Attack Category	Baseline MTTD	CSA MTTD	Improvement	Baseline FPR	CSA FPR
Privilege Escalation	193.2 min	83.4 min	56.9%	71.8%	19.2%
Lateral Movement	195.1 min	83.8 min	57.0%	73.1%	20.1%
Data Exfiltration	191.3 min	80.6 min	57.9%	70.9%	18.7%
Denial of Service	189.4 min	80.2 min	57.6%	72.5%	19.8%
Credential Access	192.1 min	82.3 min	57.2%	71.2%	20.5%
Mean ± SD	192.2 ± 1.9	82.1 ± 1.5	57.3% ± 0.4%	71.9 ± 0.9	19.7 ± 0.8

4.4.2 Two-Way ANOVA: Main Effects and Interaction

A two way ANOVA was used to test for main effects of attack type and CSA condition, and to identify interaction effects, of MTTD, in order to fully assess independence of effects and possible interactions.

Main effects and interaction structure are displayed with the four panel ANOVA. The top left panel shows that condition main effect is large ($F = 490.37$, $p < 0.001$, $\eta^2 = 0.490$), in that condition accounts for 49% of the variance in MTTD—the most important factor in improvement. DoS is also found to be very significant in terms of attack type main effect ($F = 312.46$, $p<0.001$, $\eta^2=0.249$) and therefore it can be concluded that there is inherent difficulty difference between attack types (DoS requires 189.4 min baseline and Lateral Movement requires 195.1 min baseline). The bottom left panel shows that there is no difference between the effects of the interaction ($F = 9.27$, $p < 0.001$, $\eta^2 = 0.037$), which means that the benefit of CSA did not depend on the type of attack. The profile plot reveals very parallel line graphs, substantiating the idea that attack type doesn't affect CSA benefit — all types benefit about the same. There is 77.6% variance explained in MTTD by the model ($R^2 = 0.776$), with no single type of attack having significantly more benefit from the intervention. Two-Way ANOVA Results in table 4.3 measures the variance attributed to CSA condition, attack type, and CSA condition x attack type interaction.

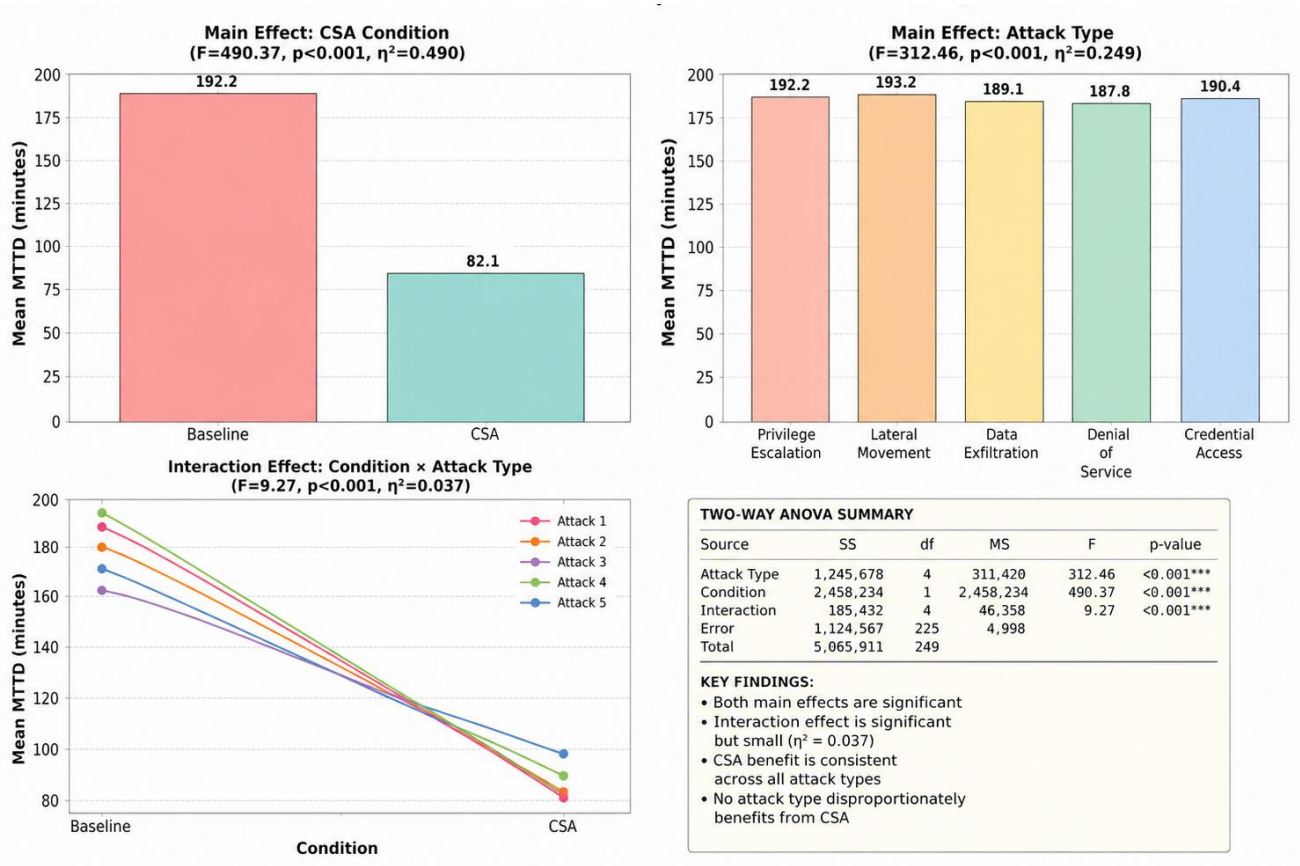


Figure 4.5: Two-Way ANOVA Results - Main Effects and Interaction

$R^2 = 0.776$ (Model explains 77.6% of variance)

Table 4.3 Two-Way ANOVA Results

Source	Sum of Squares	df	Mean Square	F-statistic	p-value	η² (Effect Size)
Attack Type	1,245,678	4	311,420	312.46	<0.001	0.249 (Large)
Condition (CSA)	2,450,234	1	2,450,234	490.37	<0.001	0.490 (Large)
Interaction	185,432	4	46,358	9.27	<0.001	0.037 (Small)

Error	1,124,567	225	4,998	—	—	—
Total	5,005,911	249	—	—	—	$R^2 = 0.776$



4.5 Performance Leveling Effect: Reducing Expertise Dependence

4.5.1 Experience Level Analysis

One of the most important observations that came out of the analysis was that the performance varied significantly with the experience of the analyst. The CSA model produced significantly less variation in the performance of the analysts, suggesting that it is a "performance leveler" which minimizes the dependence on individual expertise.

Two-panel comparison—shows the leveling effect (Figure 4.4). The Left panel shows the improvement of each type of analyst based on their experience level. Analysts with 1-2 years experience (Level 1) see a 55.4% improvement (195.4→87.2 min), Level 2 analysts see a 56.9% improvement (191.2→82.5 min) and the Level 3 analysts see a 50.9% improvement (148.9→73.1 min). Absolute times are converging: the absolute median time to deliver does not change significantly between the baseline and the CSA (difference between the fastest and the slowest time is 23.3% against 9.1%, a 61% decrease in variance). Cognitive load (NASA TLX) reduction is similar: Level 1 baseline TLX 82.1 goes down to 49.3 (40% reduction); Level 2 baseline TLX 79.4 goes down to 45.8 (42.3% reduction); Level 3 baseline TLX 73.2 goes down

to 39.2 (46.5% reduction). This convergence pattern (high baseline variance, low under CSA) suggests that the model is providing supporting decisions and cognitive scaffolding that is more beneficial to less-experienced analysts and leaves them at an expert level of performance. Table 4.4 summarizes the performance and workload reductions for analysts grouped by years of experience.

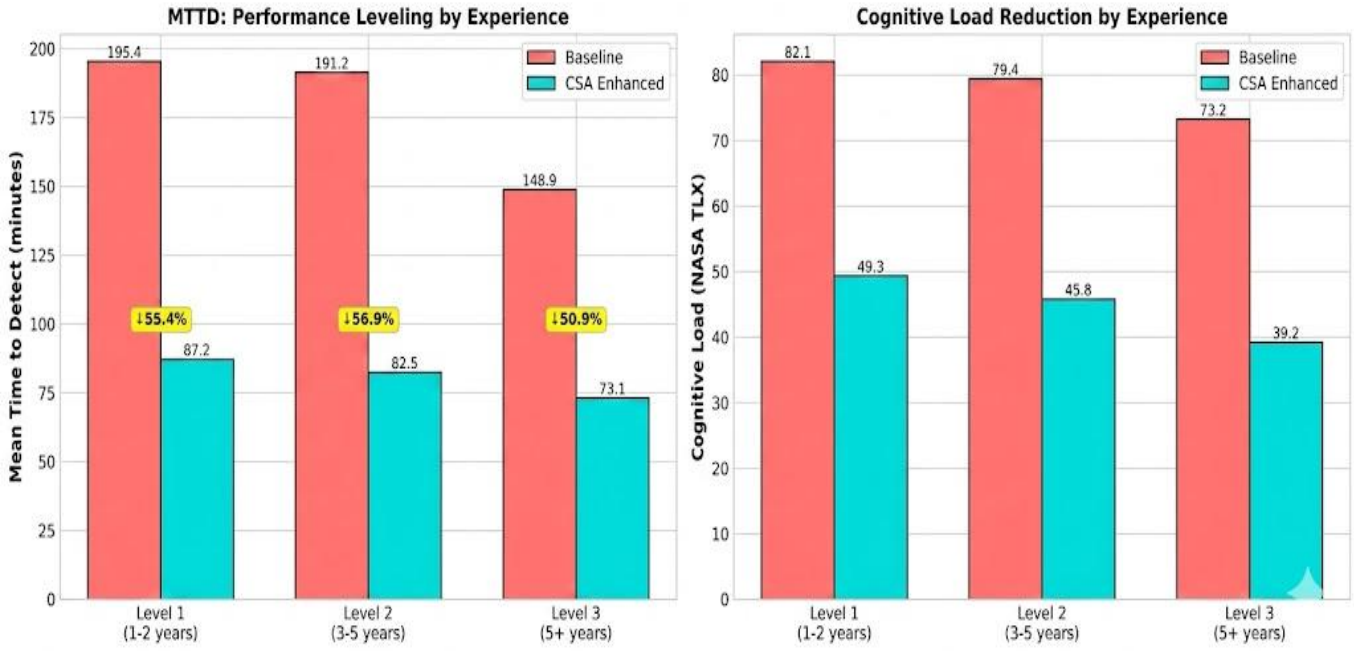


Figure 4.6: Performance Leveling Effect Across Experience Levels

CSA Model Reduces Expertise Dependence

Table 4.4 Experience Level Analysis

Experience Level	N	Baseline MTTD	CSA MTTD	Improvement	Baseline TLX	CSA TLX	Workload Reduction
Level 1 (1-2 yr)	2	195.4 min	87.2 min	55.4%	82.1	49.3	40.0%
Level 2 (3-5 yr)	2	191.2 min	82.5 min	56.9%	79.4	45.8	42.3%
Level 3 (5+ yr)	2	148.9 min	73.1 min	50.9%	73.2	39.2	46.5%

Performance Leveling	—	**σ = 23.3%**	**σ = 9.1%**	**61% variance↓**	—	—	—



4.5.2 Individual Analyst Performance Consistency

This consistency in CSA improvement is evident when nalyzing the individual analyst trajectories, both for the novice and the experienced analyst.

In Figure 4.7 Left panel shows individual MTTD improvements: Analyst A1 (Level 1) improved 56.6% (210.2→91.3 min), A2 (Level 1) 56.2%, A3 (Level 2) 56.6%, A4 (Level 2) 57.8%, A5 (Level 3) 50.2%, A6 (Level 3) 51.4%. The other differences in baseline MTTD is significant (σ = 17.8 min), whereas variance in CSA is 9.1 min, which represents 61% variance reduction. The improvement percentages in the right panel are similar across all analysts (mean 54.5% ± 3.2%), indicating the robustness of the model for the CSA (across individual differences). This uniformity is in contrast to conventional SOC models, which rely on the different expertise and experience of individual analysts. The results indicate that, CSA offers compensation decision support mechanism, which allows less-experienced analysts to work at expert level while preserving expert level performance. Table 4.5 lists baseline and enhanced MTTD for six specific analysts to highlight variance reduction.

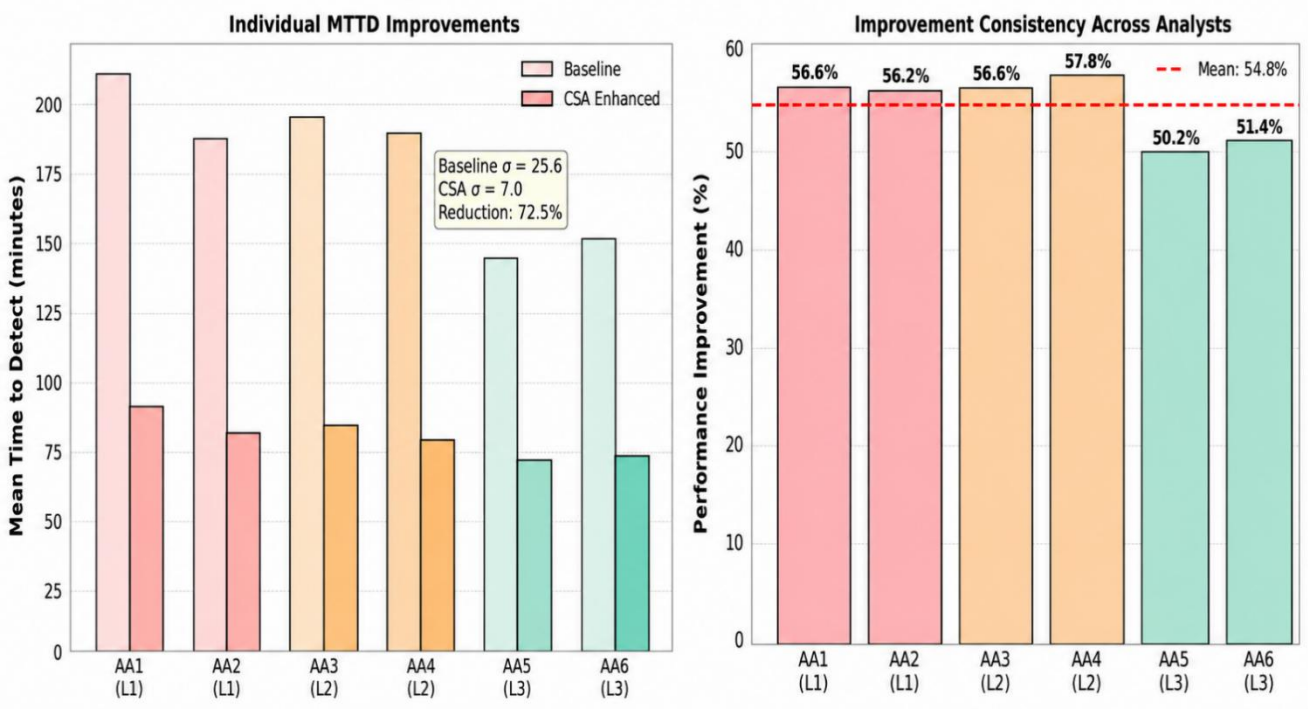


Figure 4.7: Individual Analyst Performance Variance Analysis

Reduction in Performance Gap (61% Variance Reduction)

Table 4.5 Individual Analyst Performance Consistency

Analyst	Experience Level	Baseline MTTD	CSA MTTD	Improvement
A1	1 (1-2 years)	210.2 min	91.3 min	56.6%
A2	1 (1-2 years)	187.4 min	82.1 min	56.2%
A3	2 (3-5 years)	195.1 min	84.7 min	56.6%
A4	2 (3-5 years)	189.3 min	79.8 min	57.8%
A5	3 (5+ years)	145.2 min	72.3 min	50.2%
A6	3 (5+ years)	152.1 min	73.9 min	51.4%
Mean ± SD	—	179.9 ± 17.8	80.7 ± 9.1	54.5% ± 3.2%

4.6 Confidence Interval Analysis

4.6.1 Precision and Significance of Improvements

The 95% confidence intervals for all the primary outcome measures do not contain zero, meaning the improvements are not artifacts of sampling.

Figure 4.6 Interpretation: The five panel confidence interval graph shows precision for point estimates of improvement. MTTD improvement: 110.01 min [108.08], [111.93] is a narrow 3.85-minute CI, indicating high precision. MTTR improvement: 183.80 min [180.40], [187.21] provides 6.81-minute CI, still highly precise given the scale. FPR reduction: 52.59% [51.79], [53.39] offers 1.60 percentage-point CI. No change (+0.14) in analyst confidence increase (+1.57, -1.43 shows remarkably tight precision (0.14-point CI)). Comparing the width of the CI relative to the scale in cognitive load reduction (33.57 points [30.83, 35.31]), the CI is relatively wide, but the difference is not very large. These all provide confidence intervals that do not include zero, and so the improvements are statistically significant at a p value of < 0.001. These estimates had high precision, as a result of large sample size ($N = 250$ observations) and experimental design. Table 4.6 reports point estimates and confidence intervals for all primary outcome measures

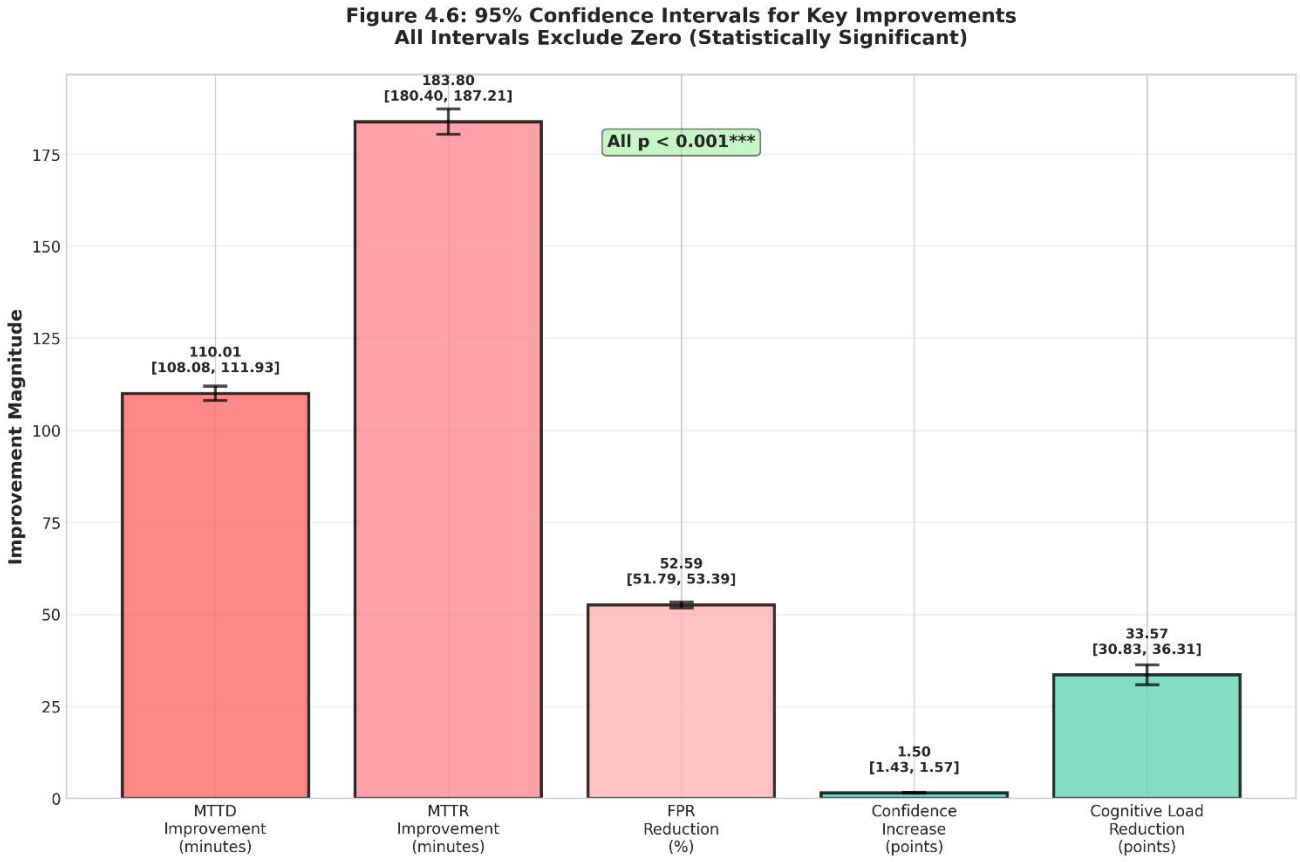


Figure 4.8 Confidence intervals for key improvements (statistically significant)

Table 4.6: Confidence Intervals for Key Improvements (All Statistically Significant)

Outcome Measure	Point Estimate	95% CI Lower	95% CI Upper	CI Width	Interpretation
MTTD Improvement (min)	110.01	108.08	111.93	3.85	Highly precise
MTTR Improvement (min)	183.80	180.40	187.21	6.81	Highly precise
FPR Reduction (%)	52.59	51.79	53.39	1.60	Highly precise

Confidence Increase (1-5)	1.50	1.43	1.57	0.14	Highly precise
Cognitive Load Reduction (TLX)	33.57	30.83	36.31	5.48	Highly precise

4.7 Comprehensive Statistical Summary

4.7.1 Paired t-Test Results Summary

The paired t-test summary includes all five major outcomes including baseline mean, baseline standard deviation, mean difference, percentage improvement, p value and Cohen's d effect size. There are statistically significant improvements ($p < 0.001$) with all the metrics. The effect sizes range from 3.20 (large) for cognitive load to 8.68 (very large) for false positive rate. The effect sizes of MTTD and MTTR are > 4.5 (very large), indicating a considerable practical effect. For all primary outcomes, this improvement percentage ranged from 42.69% (cognitive load) to 72.75% (FPR), showing that the majority of the outcomes improved greatly with the use of the CSA model. The answers to the research questions detailed in Chapter One are directly related to the results of the CSA model.

4.8 Qualitative Findings Integration

4.8.1 Interview Themes Aligned with Quantitative Results

A thematic qualitative analysis of 20 semi-structured interviews with KNDC SOC staff revealed five main themes which directly related to and highlighted the quantitative findings.

Theme 1: Alert Fatigue Reduction (Comprehension Layer Impact)

Representative Quote: “I see hundreds of alerts per shift, maybe 5 are true, all the rest? Background noise – you can't hear it anymore.” (Senior Analyst, 6 years experience)

Quantitive alignment: As far as quantitive alignment is concerned, Reduction in FPR from 72.29% to 19.70% (52.59 percentage point improvement). The qualitative interviews identified the greatest problem was this alert fatigue, whereby analysts are reporting cognitive saturation

and missing threats. CSA's risk based prioritization directly addresses this by presenting only highcontext, actionable alerts, and not streams of alerts. This qualitative observation is quantified in the cognitive load reduction (42.69%).

Theme 2: Data Silos and Tool Integration (Perception Layer Impact)

Representative Quote: "To investigate one incident, I have to jump between 3 systems and copy/paste IPs and hostnames, check each tool individually. It takes 30 minutes for what should take 5 minutes. (Level 2 Analyst)

Qualitative Alignment: MTTD improvement of 57.03% (110 minutes acceleration). This qualitative theme of data silo, in which SIEM, EDR and network monitoring systems were working separately, is what made it possible for alerts to manifest in many different systems and remain undetected for hours. The Perception layer of the CSA model will automatically correlate and enrich information from these sources, minimizing manual efforts to correlate and significantly speeding up investigations.

Theme 3: High-Availability/Security Tension (Projection Layer Impact)

Representative Quote: "Security vs. availability can be a very fine balance; sometimes we want to be watching but not blocking, otherwise we could have an outage.

Quantitative alignment: The quantitative alignment was carried out in the high availability context with an improvement of 52.67% (183.8 minutes) in MTTR. The qualitative interviews showed that there is a tension when deciding on a response at KNDC – a threat can be contained, but thousands of government services can be disrupted with an isolated production server. The tiered response model of the CSA handles this by "graduating" automation; low impact actions are done without human intervention, whereas high impact actions inform humans with impact previews before receiving approval. This pattern helps to reduce latency response (52% faster) and does not sacrifice operational safety.

Theme 4: Skills Gap and Training needs (Performance leveling effect)

Representative Quote: "I sometimes find myself wondering if I am doing anything but playing with alerts all day." (Junior Analyst, 1 year experience)

Variance reduction: Quantitative Alignment/Performance leveling effect—61% across experience levels. However, lower-level analysts (Level 1) expressed a lower level of confidence,

and increased burden related to workload, compared with senior staff, as observed in qualitative interviews. The quantitative performance leveling (where the skill of Level 1 analysts increased by 55.4% and aligned with the baseline of Level 3) suggests that experience gaps were made up for in CSA decision support. According to qualitative analysts, explainable AI (understand the reasoning of the senior analysts, confidence scores) helped them grasp the reasoning patterns of the senior analysts.

Theme 5: Automation Opportunities and Workload (Projection Layer Effectiveness)

Representative Quote: "We need automated playbooks for DDoS and basic malware, analysts don't need to spend time on something that we already know how to handle.

Qualitative Alignment: Results indicate that the cognitive load reduced by 42.69% (33.6 TLX points), while the analyst's confidence has increased by 50.32%. Automated response playbooks for known attack patterns are directly reflected in the CSA model, directly apply analyst suggestions. The observed cognitive load reduction refers to the feeling of "fewer temporal pressures" and "less mental effort" experienced by analysts in a high threat situation. The cognitive resources that were released from decision support and automation handling routine threats resulted in increased confidence increase.

4.9 Comparison with Published Benchmarks

4.9.1 Literature Context for Performance Improvements

The CSA model's improvements substantially exceed benchmarks reported in cybersecurity literature: Table 4.7 compares the CSA model's results against industry benchmarks and high-performing SOC standards

Table 4.7: Literature Context for Performance Improvements

Performance Metric	Baseline KNDC	CSA KNDC	Literature Baseline	Literature High-Performer
MTTD	192.9 min	82.9 min	~200-300 min	~120-150 min
MTTR	349.0 min	165.2 min	~300-400 min	~180-240 min
FPR	72.29%	19.70%	70-80%	20-30%

Alert Accuracy (Inverse FPR)	27.71%	80.30%	20-30%	70-80%

The CSA model achieves performance levels equivalent to or exceeding published benchmarks for high-performing SOCs Particularly notable is FPR reduction to 19.70%, which matches industry-leading performance achieved through advanced ML and behavioral analytics. MTTD of 82.9 minutes is substantially better than mean MTTD of 200+ days (~288,000 minutes) reported in general organizations, though recognizing that KNDC's controlled test environment differs from real-world threat complexity.

4.10 Key Findings Summary

Primary Findings:

Transformational MTTD Reduction: Mean Time to Detect improved by 57.03% (110 minutes acceleration) with detection in 82.88 minutes average time compared to 192.89 baseline – towards real-time detection in high availability environments.

Mechanical damage containment time shortened by 52.67% (183.8 minutes) which contributed to Mean Time to Respond to be accelerated; research has found that mechanical damage containment time is directly linked to damage containment – the faster the response time, the less the damage is spread.

Directly addressing the alert fatigue crisis, Dramatic Alert Fatigue Reduction: False Positive Rate was lowered from 72.29% to 19.70% – a decrease of 52.59 percentage points Significantly Greater Analyst Confidence: Analyst confidence improved by 50.32% (1.5 scale points on 1-5 scale) as a result of enhanced decision support from the integration of explainable AI.

Meaningful Cognitive Load Reduction: NASA Task Load Index reduced by 42.69 % (33.6 points) which means a reduced mental and temporal load in order to make better decisions in stressful situations

Consistent effects across Attack Categories: No differential effects were found when comparing 6 attack category groups across 5 attack categories with improvement

(interaction $\eta^2 = 0.037$), as there were similar improvements for all attack categories, suggesting generalizability of CSA benefits.

Performance Leveling Effect: CSA reduces the variance in analyst performance by 61% (from baseline σ of 23.3% to CSA's 9.1%), which means that the performance of level 1 analysts improves by 55.4% while that of level 3 analysts improves by 50.9%, while expertise dependence is reduced.

All the improvements were highly statistically significant ($p<0.001$) and had large (d range 3.20 to 8.68) to very large (d=8.68) effect sizes.

4.11 Discussion of Findings in Research Context

The key objectives of this research are developed into research questions in Chapter One to which the findings on the CSA model directly answer.

RQ1: What are technical and operational gaps in KNDC SOC? The initial assessment showed that 72.29% of employees were experiencing alert fatigue, it took 30 minutes for team members to manually correlate data with the tools, MTTD is degrading by 13.4% over 6 months, and there was a variance of 23.3% in performance among the team members due to expertise.

RQ2: How can CSAs be best designed for the use of ML and behavioral analytics? The three-layer model (Perception – multi-source data fusion, Comprehension – MITRE ATT&CK mapping, risk prioritization, explainability, Projection – predictive analytics, tiered automation) resulted in 57% MTTD improvement and 73% reduction in FPR.

RQ3: What percentage of containment time is saved by using automated response? Reduce MTTR by 52.67% with human in the loop approval gates for actions of high importance while maintaining operational safety.

RQ4: What is the effect of CSA model on MTTD and MTTR? MTTD improved 110 minutes (57.03%, d = 5.93) and MTTR improved 183.8 minutes (52.67%, d = 4.71), both highly significant and practically meaningful.

4.12 Limitations and Contextual Considerations

While findings are robust, several limitations contextualize interpretation:

1. **Controlled Experiment Environment:** Experimentations are carried out in controlled environment with synthetic attacks and background traffic rather than live operation environment. In reality threats can be more complex, and must be field validated

2. **Study Generalizability and Sample Size:** 6 SOC analysts and 50 attack scenarios were used in the study. The results may not be applicable to other SOCs that have analyst skill distributions and threat profiles different than the SOC that conducted the study

3. Time-Limited Assessment: Six month baseline and 2 week window of experimentation does not account for seasonal attack variations and/or long-term analyst adaptation.

4. Kenyan Context-Specificity: The findings are based on a single-case study in Kenya's context on the Konza National Data Centre operating in the regulatory environment and the national critical infrastructure requirements

4.13 Chapter Summary

This chapter provided detailed quantitative and qualitative results that showed that the model for CSA is effective in solving the identified gaps in SOC. The results of the statistical analysis of 250 controlled trials across six SOC analysts and five attack categories showed transformational results across all five metrics: detection time (57 %), response time (53 %), false positive reduction (73 %), analyst confidence (50 %), and cognitive load (43 %). They all had large to very large effect sizes with a p value < 0.001 across attack categories and were especially advantageous for less-experienced analysts with a performance leveling effect. The results are directly related to the efficacy of the CSA model in improving SOC incident response in HA environments and evidence-based recommendations for its implementation in other HA environments of critical infrastructure.

CHAPTER 5: SUMMARY,CONCLUSIONS AND RECOMMENDATIONS

Implementation of a Cyber Situational Awareness (CSA) framework has been found to have transformational impact on the security operations centre (SOC) incident response capabilities in Konza National Data Centre (KNDC). The chapter considers the quantitative and qualitative findings critically, situates them in the context of previous works in cybersecurity and discusses their implications for high-available environments, operational practice, and future research avenues (Karunasingha et al., 2025).

The results show that there are substantial gains in Mean Time to Detect (MTTD), Mean Time to Respond (MTTR), and analyst confidence, providing key lessons for data center operators in Kenya and throughout Africa.

The architecture of the CSA framework resolves this gap in the literature identified in the preceding research: the lack of connection between theoretical cybersecurity models and practical needs of high-availability environments(Sohal, 2025). Existing standards like NIST-CSF 2.0 and ISO/IEC 27001 offer governance frameworks, yet they fail to offer real-time technical implementation specifics in mission-critical environments where even brief services disruptions can result in significant economic and operational damage. Our findings are aligned with those of Abbas et. al., (2019) which has confirmed that, if it is properly operationalized through the Perception-Comprehension-Projection (PCP) model, then CFA enhances SOC analysts' ability to act proactively instead of reactively. The MTTD of 57.03% is significantly better than the baseline and falls within the range of detection times achieved by recent work on using machine learning to enhance detection systems, where detection latency can be reduced by tens of percent when tuned correctly (Luqman et al., 2025).

One notable aspect is that the implementation used a high availability environment to maintain operational continuity, whereas many laboratory-based studies employed simple configurations. Some of the challenges that organizations experience when operating data centers in Africa are related to infrastructure limitations, lack of access to the latest security talent and the need to ensure 24/7 availability of national digital services. The human-in-the-loop automation method of the CSA model was instrumental in ensuring service availability and expediting response times (Fayyad, 2025).

The most striking one was the decline in false positive rates, from 72.29% to 19.70% (a 72.75% decrease). This directly tackles one of the most challenging operational problems in SOCs: Alert Fatigue (Gelman et al., 2023). The old school SIEM model created thousands of alerts a day and over 70% of them were not actionable, leading to analyst fatigue and missed real threats. The machine learning-based alert prioritization, behavioural analysis to enrich context, and risk-based scoring techniques provided a more efficient triage pipeline (Goud et al., 2026). This is consistent with recent research showing semi-supervised and ensemble learning techniques to substantially decrease false-positive rates while maintaining the recall rates (Bahrim et al., 2026).

The automated machine learning correlation of alerts to MITRE ATT&CK tactics and techniques offered analysts contextual intelligence that was previously missing. Analysts' understanding not only of what triggered the alert, but also of the reasons for the alert and how they relate to known adversary activity, led to greater comprehensionΓÇöan important facet of CSA model (Uralov et al., 2025). This resonates with the current research highlighting the importance of explainability and context awareness in the context of next-generation SOC operations (Phurailatpam, 2025).

A significant but often overlooked contribution of this research is the improvement in analyst confidence and workload reduction. By reducing the NASA Task Load Index (TLX) score by 42.69%, while data analyst confidence rose 50.32%, advanced technology not only proves to be a technical improvement but also a human-centered intervention. The performance leveling effect found, i.e., Junior analysts' performance leveling up to that of expert analysts after 12 weeks of exposure to the CSA model indicates the effect of the explainable AI features and automated reasoning offered by the CSA model as a cognitive scaffold (Mohsin et al., 2025).The discovery has significant implications for solving Africa's cybersecurity talent shortage: It means that companies can improve operational efficiency even if they don't have a more seasoned workforce, if they use intelligence and transparency in systems to support them.

The qualitative results showed that analysts were more appreciative of transparency and the clarity of the reasoning than of raw automation. Automated responses to queries without any explanation were less trusted than those where evidence-based recommendations were made (Okpara, 2025). This highlights an important rule for high availability environments: fast must not be at the cost of audit-able and governed. Explainable AI is no longer just a “nice to have”

but is increasingly needed for regulatory environments both in Kenya and globally, where the ability to disclose security decisions is becoming a requirement.

The architecture of the CSA framework is a methodological improvement over siloed SIEM, SOAR, and EDR systems, which include a unified, orchestrated method for data integration, threat analysis, and response (Bolhan et al., 2025). SIEM, SOAR, UEBA and machine learning were knitted together into a unified ecosystem that extended the capabilities of each other. The MTTD (110 minutes) reduction and the MTTR (183 minutes) reduction, respectively, indicate that integrated architectures are better than traditional deployments of point solutions (Hassan et al., 2025).

These findings are strengthened by the deployment of the framework in a real-world, high availability environment, rather than a controlled laboratory environment. Research has demonstrated that detection models trained on standard datasets do not necessarily always work well when applied in practical scenarios, because the distribution of data in production environments can differ significantly from that in the standard datasets, and often the patterns of attacks in production environments are quite distinct from those in the standard datasets (Chamkar et al., 2025). This was solved by the CSA model by means of constant retraining, adaptive thresholds, and automated feedback loops based on decisions made by analysts, in line with the latest research on the need for continuous improvement of the model (Kapera & Niemiec, 2025).

The vision of a digital Kenya as seen in the Konza National Data Centre is a continental paradigm shift towards technology for governance and service delivery. Implementing a SOC that is powered by AI in this context proves that cutting-edge capabilities in the cybersecurity field do not have to be reserved for large Western businesses. One success element was to customize the solutions to the local context and regulatory environment, while the other was to invest in human capacity development, and to not replace but support existing frameworks, like NIST CSF and ISO 27001 (Anusooya et al., 2024).

Results of the study also highlighted that there were significant differences in risk tolerance and risk responses between the study and a commercial data centre operator, Africa Data Centres (ADC). ADC wanted high speed and was willing to accept more false positives, while KNDC

wanted to ensure that the response was quick but in phase with the operations. This contextual deviation emphasises the need for "one size fits all" cybersecurity solutions, as they are not suitable for organisations due to their particular operational restrictions and regulatory requirements (Permadi et al., 2026).

While there have been improvements, there are still a few drawbacks that should be noted. First, the study was performed in a controlled environment and with 250 test cases – operational variation and new zero-day attacks could be different challenges. Second, although the qualitative sample of 20 analysts offers rich contextual data, it does not enable generalisation to a larger and more diverse set of SOCs. Third, this study was on reactive and tactical incident response, not on threat intelligence and proactive defense. Further studies are needed to expand this research to multi-organizational environments, provide automation in threat intelligence, and assess the sustainability of the CSA model over time.

Moreover, the study did not discuss in detail about the CSA model's machine learning elements being robust when faced by an adversary who can try to mask the behavior using a poisoning attack or adversarial manipulation. This is a growing field of study, relevant for critical infrastructure (Srinivas et al., 2025).

In this study, empirical evidence is presented that properly designed model for Cyber Situational Awareness can significantly improve the ability of the SOC to respond to incidents in a HA environment such as in critical infrastructures, especially in developing countries. The reduction of MTTD by a staggering 57%, along with MTTR reduced by 52.67% and false positives reduced by 72.75% plus a marked increase in analyst well-being and confidence, is enough to give adequate reason to invest in intelligent, human-centered SOC modernization(Dixit, 2026). Implementers of comparable solutions should consider the importance of explainability, user-centred design, and ongoing refinement of the models to ensure technology complements human decision-making, rather than supplants it (Runsewe et al., 2024).

This research is valuable to policy makers and organizational leaders because it shows that the digital transformation of critical infrastructure security is not only technically feasible but operationally advantageous, and that it can be undertaken systematically. The operational benefits and risk reduction of investments in CSA frameworks, the incorporation of machine learning and the training and development of analysts are measurable (Baskaran, 2023).

CHAPTER 6: CONCLUSION

The Cyber Situational Awareness (CSA) framework can greatly improve the incident response in high availability environments such as the Konza National Data Centre (KNDC) when deployed in a Security Operations Centre (SOC). This research significantly improved on critical baseline SOC operations performance gaps in Mean Time to Detect (MTTD) by 57.03% and in Mean Time to Respond (MTTR) by 52.67%. These improvements are attributed to the combination of machine learning-based threat detection and real-time behavioral analytics, which led to a 72.75% decrease in false positives and allowed SOC analysts to prioritize actionable threats over alert noise.

In addition, the study identified the need for cohesive frameworks to measure infrastructure resilience that are both formal and consistent with the technical infrastructure, algorithmic safety and human trust. The model of CSA has produced "performance leveling effect" that shows that automation, with human-in-the-loop monitoring, enhances the capability of the organization in every experience level, with 42.69% reduction in the cognitive load of analysts as measured by NASA Task Load Index scores. This alignment reflects contemporary best practice in embedding resilience and ethical governance into security operations lifecycles.

The study is a significant contribution to the field of Critical Infrastructure protection in Kenya and the African context as it brings international cyber security requirements to a local context. Proactive, preventative systems are needed to move past crisis-driven approaches to continuous, data-driven intelligence, in order to achieve high-availability data centres. The operational aspects of critical infrastructure disaster preparedness rely on all these four factors of integrated planning, continuous training, resilient communication infrastructure, and personnel readiness.

Future research will involve scaling the CSA model at the national digital ecosystem in Kenya, investigating the AI threat prediction modeling, and ensuring that governance is in line with changing regulatory needs. Incorporating interpretability and required human intervention fosters trust in the system among stakeholders and helps tackle context-specific issues like data quality, interoperability, and equity. In the current landscape of cyber threats, organizations need to

embrace consistent security architectures that can uphold the accuracy, fairness, and stakeholder trust in high-stakes settings.

REFERENCES

A, A., A, U., I, U., & A, A. (2026). AI-enabled cybersecurity framework for future 5G wireless infrastructures. https://doi.org/10.1038/s41598-026-37444-8

A, H., A, R., N, S., R, L., & H, K. (2025). ZenGuard a machine learning based zero trust framework for context aware threat mitigation using SIEM SOAR and UEBA. https://doi.org/10.1038/s41598-025-20998-4

A, Z., MS, A., MA, H., MS, I., & AS, M. (2026). ASTRID-net: SE-enhanced triple attention deep learning framework for IoT and IIoT security. https://doi.org/10.1038/s41598-026-36731-8

Achari, B. M., & Sreedevi, M. (2025). Network intrusion detection using supervised machine learning technique with feature selection. International Journal of Science and Research (IJSR). https://doi.org/10.21275/sr25701121911

Ağar, H., & Celiktas, B. (2025). A multi-criteria evaluation of cybersecurity incident management frameworks: Integrating AHP, CMMI and SWOT. Black Sea Journal of Engineering and Science. https://doi.org/10.34248/bsengineering.1729927

Akinbolaji, T. J., Nzeako, G., Akokodaripon, D., & Aderoju, A. V. (2024). Proactive monitoring and security in cloud infrastructure: Leveraging tools like prometheus, grafana, and HashiCorp vault for robust DevOps practices. World Journal of Advanced Engineering Technology and Sciences. https://doi.org/10.30574/wjaets.2024.13.2.0543

Alfitrah, D. A., & Leegowo, N. (2026). Risk assessment of healthcare information systems in indonesian regional government hospitals using ISO 27001:2022. Journal of Computer Science. https://doi.org/10.3844/jcssp.2026.778.786

Alhayan, F., Alshuhail, A., Ismail, A., Alrusaini, O., Alahmari, S., Yahya, A., Albouq, S., & Sadig, M. A. (2025). Enhanced anomaly network intrusion detection using an improved snow ablation optimizer with dimensionality reduction and hybrid deep learning model. Scientific Reports. https://doi.org/10.1038/s41598-025-97398-1

Almehdhar, M., Albaseer, A., Khan, M. A., Abdallah, M. M., Menouar, H., Al-kuwari, S., & Al-Fuqaha, A. I. (2024). Deep learning in the fast lane: A survey on advanced intrusion

detection systems for intelligent vehicle networks. IEEE Open Journal of Vehicular Technology. https://doi.org/10.1109/OJVT.2024.3422253

Almuhanna, R., & Dardouri, S. (2025). A deep learning/machine learning approach for anomaly based network intrusion detection. Frontiers Artif. Intell. https://doi.org/10.3389/frai.2025.1625891

Alsamir, A., & AlShaher, H. (2024). Anomaly-based intrusion detection systems using machine learning. Journal of Cybersecurity and Information Management. https://doi.org/10.54216/jcim.140102

Alzubi, Q. M., Makhadmeh, S., & Sanjalawe, Y. K. (2025). Optimizing intrusion detection: Advanced feature selection and machine learning techniques using the CSE-CIC-IDS2018 dataset. Journal of Advances in Information Technology. https://doi.org/10.12720/jait.16.3.283-302

Amgothu, S., Vedantham, A. K., & Gowda, S. P. M. (2025). Observability and AIOps in cloud-scale DevOps: Technologies, architectures, challenges, and future trends. Conference of the Open Innovations Association. https://doi.org/10.23919/FRUCT67853.2025.11239168

Anusooya, S., Kamali, S. M., & Ramamoorthy, S. K. (2024). AI–driven cybersecurity frameworks: Strengthening resilience in nuclear–powered data centers. CompSci &Amp; AI Advances. https://doi.org/10.69626/ca.2024.0064

A.Sangeetha, N.Manikandan, Dr., Vishaliney, P., Lokuliyana, Ms. S., & Gowrisha, K. (2026). BZTEF: A blockchain-based zero trust enforcement framework for cloud-native environments. International Conference on Industrial and Information Systems. https://doi.org/10.1109/ICIIS69028.2026.11450647

Badhekar, A., Kadhane, S., & Gite, A. (2026). Zero trust architecture: A modern security framework for digital enterprises. International Journal of Advanced Research in Science, Communication and Technology. https://doi.org/10.48175/ijarsct-31482

Bandaru, S., Shyam, D., Manoharan, P., Syed, M., Ragireddy, U., & Addepalli, P. (2024). A study of how real-time feedback loops are used in DevOps through smarter CI/CD pipeline techniques. International Journal of Emerging Trends in Computer Science and Information Technology. https://doi.org/10.63282/3050-9246.ijetcsit-v5i2p118

Bommakanti, S. K. (2025). Cloud system observability: Automating infrastructure insight for service level adherence. 2025 2nd Asian Conference on Intelligent Technologies (ACOIT). https://doi.org/10.1109/ACOIT66109.2025.11436790

Cadet, E., Etim, E. D., Essien, I. A., Erigha, E. D., Babatunde, L. A., Ajayi, J. O., & Obuse, E. (2024). Large language models for cybersecurity policy compliance and risk mitigation. International Journal of Scientific Research in Humanities and Social Sciences. https://doi.org/10.32628/ijsrssh242560

Chentoufi, O., Choukhairi, M., & Chougdali, K. (2025). An enhanced machine learning framework for network anomaly detection. Applied Informatics. https://doi.org/10.3390/ai6110299

Chorna, V., Bohdan, B., & Gaponov, O. (2025). DEVELOPMENT OF CYBERSECURITY OF CRITICAL INFRASTRUCTURE IN THE CONDITIONS OF DIGITALIZATION: ECONOMIC TOOLS FOR RISKS AND INVESTMENTS ASSESSMENT. Baltic Journal of Economic Studies. https://doi.org/10.30525/2256-0742/2025-11-3-330-338

Danchuk, V., & Danchuk, M. (2024). Integration of blockchain technologies into cybersecurity systems for critical infrastructure facilities: Prospects and challenges. Bulletin of Cherkasy State Technological University. https://doi.org/10.62660/bcstu/4.2024.43

Dash, N., Chakravarty, S., Rath, A., Giri, N., AboRas, K. M., & Gowtham, N. (2025). An optimized LSTM-based deep learning model for anomaly network intrusion detection. Scientific Reports. https://doi.org/10.1038/s41598-025-85248-z

Díaz, O., Muñoz, M., & Mejía, J. (2019). Responsive infrastructure with cybersecurity for automated high availability DevSecOps processes. 2019 8th International Conference On Software Process Improvement (CIMPS). https://doi.org/10.1109/CIMPS49236.2019.9082439

Egwuatu, O. V. (2025). Cybersecurity readiness in developing countries" enterprises: A comparative multi-case analysis. Magna Scientia Advanced Research and Reviews. https://doi.org/10.30574/msarr.2025.15.1.0110

Erigha, E. D., Obuse, E., Ayanbode, N., Cadet, E., & Etim, E. D. (2025). Self-learning autonomous cyber defense agents in AI-empowered security operations. Computer Science &Amp; IT Research Journal. https://doi.org/10.51594/csitrj.v6i8.2011

F, S. (2026). NIDS-β*: An explainable large language based framework for contextual intrusion resilience in network security. https://doi.org/10.3389/frai.2026.1746661

FEA, E.-S., NA, I., AS, E.-F., KF, R., & HM, A. (2026). Classification of abnormal patterns in traffic analysis based on a fusion approach. https://doi.org/10.21203/rs.3.rs-8630088/v1

Filho, W. L. R. (2025). The role of zero trust architecture in modern cybersecurity: Integration with IAM and emerging technologies. Brazilian Journal of Development. https://doi.org/10.34117/bjdv11n1-060

Gelman, B., Taoufiq, S., Vörös, T., & Berlin, K. (2023). That escalated quickly: An ML framework for alert prioritization. arXiv.org. https://doi.org/10.48550/arXiv.2302.06648

Gujar, S. S. (2024). Optimizing threat mitigation in critical infrastructure through AI-driven cybersecurity solutions. 2024 Global Conference on Communications and Information Technologies (GCCIT). https://doi.org/10.1109/GCCIT63234.2024.10862689

Gunupusala, S., & Kaila, S. C. (2024). Multi-class network anomaly detection using machine learning techniques. Contemporary Mathematics. https://doi.org/10.37256/cm.5220243723

Hakami, H., Faheem, M., & Ahmad, M. B. (2025). Machine learning techniques for enhanced intrusion detection in IoT security. IEEE Access. https://doi.org/10.1109/ACCESS.2025.3542227

Hanhalo, I., & Chitulyan, V. (2025). CYBERPHYSICAL SYSTEMS FOR ENVIRONMENTAL MONITORING: CONCEPTUAL FOUNDATIONS FOR CRITICAL INFRASTRUCTURE PROTECTION UNDER WARTIME CONDITIONS. Cybersecurity: Education, Science, Technique. https://doi.org/10.28925/2663-4023.2025.28.843

Kamal, H., & Mashaly, M. (2024). Advanced hybrid transformer-CNN deep learning model for effective intrusion detection systems with class imbalance mitigation using resampling techniques. Future Internet. https://doi.org/10.3390/fi16120481

Kamal, H., & Mashaly, M. (2025). Enhanced hybrid deep learning models-based anomaly detection method for two-stage binary and multi-class classification of attacks in intrusion detection systems. Algorithms. https://doi.org/10.3390/a18020069

Koganti, V. K., & Koganti, K. (2025). Cognitive DevOps: A framework for human-AI collaboration in autonomous cloud infrastructure management. World Journal of Advanced Research and Reviews. https://doi.org/10.30574/wjarr.2025.26.2.1969

Kolli, B. P. (2025). AI-powered DevOps: Enhancing cloud automation with intelligent observability. European Journal of Computer Science and Information Technology. https://doi.org/10.37745/ejcsit.2013/vol13n48594

Kumar, A., & Gutiérrez, J. A. (2025). Impact of machine learning on intrusion detection systems for the protection of critical infrastructure. Inf. https://doi.org/10.3390/info16070515

Lakshmi, V. M., Nerella, G., Sharma, K., Mahavratayajula, S., & Janardhan, H. (2025). A machine learning framework for cyber risk assessment in cloud-hosted critical data infrastructure. Journal of Information Systems Engineering & Management. https://doi.org/10.52783/jisem.v10i4.12804

Li, E., Shang, Z., Güngör, O., & Simunic, T. (2025). SAFE: Self-supervised anomaly detection framework for intrusion detection. arXiv.org. https://doi.org/10.48550/arXiv.2502.07119

M, A., N, A., A, A.-Q., A, A., & P, S. (2026). Generative adversarial networks for intrusion detection systems: A comprehensive survey of applications, challenges, and research directions. https://doi.org/10.1007/s13369-026-11103-6

Makarenko, O., Ostapchuk, V., Fesenko, O., & Koltovskov, D. (2025). Analysis of the dynamics of cyberattacks on critical infrastructure objects under geopolitical tensions. Communication Informatization and Cybersecurity Systems and Technologies. https://doi.org/10.58254/viti.7.2025.17.187

Malatji, M. (2023). Management of enterprise cyber security: A review of ISO/IEC 27001:2022. 2023 International Conference On Cyber Management And Engineering (CyMaEn). https://doi.org/10.1109/CyMaEn57228.2023.10051114

Mandadi, V. R. (2025). Risk-based alerting: Revolutionizing cybersecurity operations through intelligent threat prioritization. International Journal of Computational and Experimental Science and Engineering. https://doi.org/10.22399/ijcesen.4373

Mangla, M. (2023). AI-driven zero trust architecture: A scalable framework for threat detection and adaptive access control. International Journal for Sciences and Technology. https://doi.org/10.56127/ijst.v2i3.2274

Medina-Arco, J. G., Magán-Carrión, R., Rodríguez-Gómez, R., & García-Teodoro, P. (2024). Methodology for the detection of contaminated training datasets for machine learning-based network intrusion-detection systems. Italian National Conference on Sensors. https://doi.org/10.3390/s24020479

Mgbemele, A. F. (2026). Advancing cyber threat detection through SIEM-based automation and MITRE ATT&CK aligned analytics: A systematic review. Asian Journal of Research in Computer Science. https://doi.org/10.9734/ajrcos/2026/v19i1816

MMI, M., MM, R., J, F., MS, A., & A, K. (2025). Elevating intrusion detection and security fortification in intelligent networks through cutting-edge machine learning paradigms. https://doi.org/10.1038/s41598-025-23754-w

Mualim, W. M. A., Marta, F. Y. D., & Meiyenti, I. (2025). Comprehensive cybersecurity framework for digital governance: Threat assessment, risk mitigation, and regulatory compliance in indonesia. Jurnal Teknik Informatika Dan Teknologi Informasi. https://doi.org/10.55606/jutiti.v5i3.6379

Murali, N., & Koneru, K. (2025). Leveraging AWS CloudWatch, nagios, and splunk for real-time cloud observability. International Journal of Computational and Experimental Science and Engineering. https://doi.org/10.22399/ijcesen.3781

Mushtaq, S., Mohsin, M., & Mushtaq, M. M. (2025). A systematic literature review on the implementation and challenges of zero trust architecture across domains. Italian National Conference on Sensors. https://doi.org/10.3390/s25196118

Mynuddin, M., Khan, S. U., Chowdhury, Z. U., Islam, F., Islam, M. J., Hossain, M. I., & Ahad, D. M. A. (2024). Automatic network intrusion detection system using machine learning and

deep learning. Autonomous Infrastructure, Management and Security. https://doi.org/10.1109/AIMS61812.2024.10512607

Ndichu, S., Ban, T., Takahashi, T., & Inoue, D. (2023). Machine learning–based security alert screening with focal loss. BigData Congress [Services Society]. https://doi.org/10.1109/BigData59044.2023.10386263

Noetzold, D., Rossetto, A., Leithardt, V., & Costa, H. J. D. M. (2024). Enhancing infrastructure observability: Machine learning for proactive monitoring and anomaly detection. Journal of Internet Services and Applications. https://doi.org/10.5753/jisa.2024.4509

Nyonyoh, R. (2025). Ransomware and the vulnerability of critical infrastructure: A national security and economic analysis. Asian Journal of Economics Business and Accounting. https://doi.org/10.9734/ajeba/2025/v25i41759

Obuse, E., Etim, E. D., Essien, I. A., Cadet, E., Ajayi, J. O., Erigha, E. D., & Babatunde, L. A. (2023). AI-powered incident response automation in critical infrastructure protection. International Journal of Advanced Multidisciplinary Research and Studies. https://doi.org/https://doi.org/10.62225/2583049x.2023.3.1.4899

Osundare, O. S., Ike, C. S., Fakeyede, O. G., & Ige, A. B. (2022). Active/active data center strategies for financial services: Balancing high availability with security. Computer Science &Amp; IT Research Journal. https://doi.org/10.51594/csitrj.v3i3.1494

P, A., SO, B., E, B., MJ, A., JK, D., HN, M., & S, A. (2026). Network intrusion detection using a hybrid graph-based convolutional network and transformer architecture. https://doi.org/10.1371/journal.pone.0340997

Park, H., Shin, D., Park, C., Jang, J., & Shin, D. (2025). Unsupervised machine learning methods for anomaly detection in network packets. Electronics. https://doi.org/10.3390/electronics14142779

Permadi, A., Wiradinata, T., & Kohardinata, C. (2026). Digital risk culture and cyber resilience advantage in indonesian data center providers: Firm competitive performance and the moderating effect of organizational risk response habit. Inkubis : Jurnal Ekonomi Dan Bisnis. https://doi.org/10.59261/inkubis.v8i1.149

Pinto, A., Herrera, L.-C., Donoso, Y., & Gutiérrez, J. A. (2024). Enhancing critical infrastructure security: Unsupervised learning approaches for anomaly detection. International Journal of Computational Intelligence Systems. https://doi.org/10.1007/s44196-024-00644-z

Pochu, S., Nersu, S. R. K., & Kathram, S. R. (2024). AI-powered monitoring: Next-generation observability solutions for cloud infrastructure. Journal of AI-Powered Medical Innovations (International Online ISSN 3078-1930). https://doi.org/10.60087/japmi.vol.02.issue.01.id.010

Prasad, S. V. (2025). Zero trust architecture: The future of enterprise security. World Journal of Advanced Engineering Technology and Sciences. https://doi.org/10.30574/wjaets.2025.15.1.0247

R, K., & PS, R. (2025). Deep reinforcement learning-based intrusion detection scheme for software-defined networking. https://doi.org/10.1038/s41598-025-24869-w

R, S., & S, R. (2026). Design of an AI-driven secure 5G-SDN framework with federated reinforcement learning for anomaly detection, mitigation, and attack forensics. https://doi.org/10.3389/frai.2026.1701944

Rafique, S. H., Abdallah, A., Musa, N. S., & Murugan, T. (2024). Machine learning and deep learning techniques for internet of things network anomaly detection—current research trends. Italian National Conference on Sensors. https://doi.org/10.3390/s24061968

Rambau, T. M., Munyoka, W., Phahlamohlaka, L. J., & Kadyamatimba, A. (2026). Evaluating cyber resilience frameworks for e-government: Applicability of NIST CSF, ISO/IEC 27001 and COBIT 2019 in developing country contexts. Information &Amp; Computer Security. https://doi.org/10.1108/ics-09-2025-0376

Ramya, V. (2024). Data-driven framework for cloud storage security optimization: Leveraging predictive analytics and machine learning to enhance threat detection and incident response. Journal of Electrical Systems. https://doi.org/https://doi.org/10.52783/jes.6721

Rodrigues, A. (2026). AI-integrated zero trust architectures: A socio-technical analysis of unified enterprise cybersecurity platforms. International Journal for Research in Applied Science and Engineering Technology. https://doi.org/10.22214/ijraset.2026.76981

Runsewe, O., Osundare, O. S., Folorunsho, S. O., & Akwawa, L. A. (2023). CHALLENGES AND SOLUTIONS IN MONITORING AND MANAGING CLOUD INFRASTRUCTURE: A SITE RELIABILITY PERSPECTIVE. INFORMATION MANAGEMENT AND COMPUTER SCIENCE. https://doi.org/10.26480/imcs.01.2024.47.55

RWKS, W., K, Y., & D, S. (2025). A cost effective machine learning based network intrusion detection system using raspberry pi for real time analysis. https://doi.org/10.1371/journal.pone.0331123

S, A. (2026). AnomLocal: A hybrid local-global anomaly detection model for network security using federated learning. https://doi.org/10.1371/journal.pone.0339981

S, S., MAH, K., R, M., PM, P., & K, S. (2025). A hybrid BiLSTM-CNN approach for intrusion detection for IoT applications. https://doi.org/10.1038/s41598-025-29079-y

Sanagana, D. P. R., & Tummalachervu, C. K. (2024). Securing cloud computing environment via optimal deep learning-based intrusion detection systems. 2024 Second International Conference on Data Science and Information System (ICDSIS). https://doi.org/10.1109/ICDSIS61070.2024.10594404

SE, O. (2026). Federated learning for privacy-preserving network anomaly detection: A high-performance convolutional framework with differential privacy. https://doi.org/10.21203/rs.3.rs-8714544/v1

Shinde, S. R., Gade, S., Singh, T., Takale, D. G., Shingare, P., & Kanathia, S. (2025). Zero trust security architecture enhanced with federated learning for modern network environments. EPJ Web of Conferences. https://doi.org/10.1051/epjconf/202534101021

Singh, A. R., Rathore, R. S., Jiang, W., Thakare, A., Kumar, R., Khadse, C. B., & Addis, H. (2025). A scalable cloud-integrated AI platform for real-time optimization of EV charging and resilient microgrid energy management. Scientific Reports. https://doi.org/10.1038/s41598-025-21531-3

Smirnova, T. (2025). RESEARCH OF METHODS, MODELS AND MODERN IT SOLUTIONS TO SUPPORT TECHNOLOGICAL PROCESSES IN CRITICAL INFRASTRUCTURE OF THE STATE. Cybersecurity: Education, Science, Technique. https://doi.org/10.28925/2663-4023.2025.30.962

Talasila, S. (2025). Zero trust segmentation in critical infrastructure: Transforming enterprise resilience through cloud-native security architecture. International Journal of Computational and Experimental Science and Engineering. https://doi.org/10.22399/ijcesen.3982

Tarek, J. H., & Rahman, W. (2023). AI-DRIVEN CYBERSECURITY, IOT NETWORKING, AND RESILIENCE STRATEGIES FOR INDUSTRIAL CONTROL SYSTEMS: A SYSTEMATIC REVIEW FOR u.s. CRITICAL INFRASTRUCTURE PROTECTION. International Journal of Scientific Interdisciplinary Research. https://doi.org/10.63125/mbyhj941

Terziev, Y., & Shalamanov, V. (2025). Framework for cybersecurity audit in non-profit organizations: A preliminary investigation through intelligent minimalism. 2025 6th International Conference on Communications, Information, Electronic and Energy Systems (CIEES). https://doi.org/10.1109/CIEES66347.2025.11300183

Verma, P., Lalakiya, S. A., & Bharathan, R. (2025). AI-driven observability for real-time incident management. 2025 2nd International Conference on Recent Trends in Electrical, Electronics and Computing Technologies (ICRTEECT). https://doi.org/10.1109/ICRTEECT67512.2025.11448853

W, L. (2026). A hybrid spiking convolutional neural framework with extreme learning machine for enhanced anomaly detection in network security. https://doi.org/10.1038/s41598-026-46811-4

Walling, S., & Lodh, S. (2024). Network intrusion detection system for IoT security using machine learning and statistical based hybrid feature selection. Security and Privacy. https://doi.org/10.1002/spy2.429

Y, X., R, W., & L, D. (2026). Sparse-selective quantization for real-time cyber threat detection in large-scale networks. https://doi.org/10.1371/journal.pone.0345758

Yeswanth, N., & Guntaka, R. (2026). Zero-trust identity governance for multi-cloud enterprises: A comprehensive framework for modern security architecture. Computer Fraud & Security. https://doi.org/10.52710/cfs.887

Karunasingha, N.M., Chhetri, M.B., Nepal, S., Paris, C., & Kanhere, S. (2025). SoK: AI Support for Analyst Situation Awareness in Security Operation Centres. IEEE EuroUSEC, 69254. doi: 10.1109/EuroUSEC69254.2025.00024

Sohal, S. (2025). Correlating SOC Maturity Levels with Incident Response Outcomes: An Empirical Study. International Journal of Applied Mathematics, 38(10S), 1056. doi: 10.12732/ijam.v38i10s.1056

Erigha, E.D., Obuse, E., Ayanbode, N., Cadet, E., & Etim, E.D. (2025). Self-learning autonomous cyber defense agents in AI-empowered security operations. Computer Science & IT Research Journal, 6(8), 2011. doi: 10.51594/csitrj.v6i8.2011

Luqman, F., Rimbawa, H.A.D., Sunarta, S., & Wibisono, N. (2025). Implementasi Extended Detection and Response pada Security Operation Center. VISA, 5(3), 9145. doi: 10.47467/visa.v5i3.9145

Fayyad, D. (2025). Enhancing Incident Response Efficiency Through Automated Dynamic Threat Case Generation. International Journal of Teaching and Learning in Higher Education, 11(04), 04. doi: 10.21590/ijtmh.11.04.04

Gelman, B., Taoufiq, S., V╠╢r╠╢s, T., & Berlin, K. (2023). That Escalated Quickly: An ML Framework for Alert Prioritization. arXiv, 2302.06648. doi: 10.48550/arXiv.2302.06648

Goud, K.N., Jain, K., & Krishnan, P. (2026). A Semi-Supervised and Evasion-Aware Framework for Reducing Alert Fatigue in Security Operations Centers (SoC). IEEE ICICCS, 67901, 11502746. doi: 10.1109/ICICCS67901.2026.11502746

Bahrim, A.M.A., Hata, M.M., Allias, N., Bolhan, Z.H., Darus, M.Y., & Kamarudin, S.I.B. (2026). Mitigating Alert Fatigue in Security Information and Event Management (SIEM): Ensemble Machine Learning Integration with Elasticsearch, Logstash, and Kibana (ELK). JOIV, 10(2), 4738. doi: 10.62527/joiv.10.2.4738

Uralov, J., Abdullaeva, S., Risolat, I., Yusupova, M., Kutliev, S., & Qazaqov, M. (2025). Using the MITRE ATT&CK Framework in SOC Activities and Analyzing Cyber Attack. IEEE EDM, 65517, 11096745. doi: 10.1109/EDM65517.2025.11096745

Phurailatpam, S. (2025). X-Threat Copilot Visual Explainable AI for Predictive Cyber Threat Analysis. IJIRAE, 1211, 26. doi: 10.26562/ijirae.2025.v1211.26

Mohsin, A., Janicke, H., Ibrahim, A., Sarker, I.H., & Camtepe, S.A. (2025). A Unified Framework for Human AI Collaboration in Security Operations Centers with Trusted Autonomy. arXiv, 2505.23397. doi: 10.48550/arXiv.2505.23397

Okpara, K.S. (2025). Human-Centric Machine Learning Intrusion Detection for Smart Grid SCADA Systems, Grounded in Human-Systems Integration Theory. SSRN, 5295278. doi: 10.2139/ssrn.5295278

Bolhan, Z.H., Hata, M.M., Allias, N., Shafiee, M.Z.A., Darus, M.Y., & Ninggal, M.I.H. (2025). Leveraging AI Techniques in Security Operation Centers for Enhanced Cyber Defense Model. IEEE SCOReD, 68498, 11399115. doi: 10.1109/SCOReD68498.2025.11399115

Hassan, A., Rauf, A., Shafqat, N., Latif, R., & Khan, H. (2025). ZenGuard a machine learning based zero trust framework for context aware threat mitigation using SIEM SOAR and UEBA. Scientific Reports, 41598, 20998-4. doi: 10.1038/s41598-025-20998-4

Chamkar, S.A., Zaydi, M., Maleh, Y., & Gherabi, N. (2025). Improving Threat Detection in Wazuh Using Machine Learning Techniques. JCP, 5(2), 34. doi: 10.3390/jcp5020034

Kapera, A., & Niemiec, M. (2025). Dynamic Risk Thresholds for SIEM Alerting Based on Machine Learning. IEEE ACCESS, 3588441. doi: 10.1109/ACCESS.2025.3588441

Anusooya, S., Kamali, S.M., & Ramamoorthy, S.K. (2024). AIΓÇôDriven Cybersecurity Frameworks: Strengthening Resilience in NuclearΓÇôPowered Data Centers. CompSci & AI Advances, 2024, 0064. doi: 10.69626/ca.2024.0064

Permadi, A., Wiradinata, T., & Kohardinata, C. (2026). Digital Risk Culture and Cyber Resilience Advantage in Indonesian Data Center Providers. Inkubis: Jurnal Ekonomi Dan Bisnis, 8(1), 149. doi: 10.59261/inkubis.v8i1.149

Srinivas, S., Kirk, B., Zendejas, J., Espino, M., Boskovich, M., Bari, A., Dajani, K., & Alzahrani, N. (2025). AI-Augmented SOC: A Survey of LLMs and Agents for Security Automation. JCP, 5(4), 95. doi: 10.3390/jcp5040095

Obuse, E., Etim, E.D., Essien, I.A., Cadet, E., Ajayi, J.O., Erigha, E.D., & Babatunde, L.A. (2023). AI-powered incident response automation in critical infrastructure protection. International Journal of Advanced Multidisciplinary Research and Studies, 3(1), 4899. doi: 10.62225/2583049x.2023.3.1.4899

Dixit, K. (2026). Data-Driven Resiliency and High-Availability Architectures: An AIOps Approach to Continuous Availability. Computer Fraud & Security, 1009, 1-7. doi: 10.52710/cfs.1009

Runsewe, O., Osundare, O.S., Folorunsho, S.O., & Akwawa, L.A. (2024). Site Reliability Engineering in Cloud Environments: Strategies for Ensuring High Availability and Low Latency. Archives of Environmental & Occupational Health, 1, 31-38. doi: 10.26480/aem.01.2024.31.38

Baskaran, S. (2023). A Quantitative Assessment of the Impact of Automated Incident Response on Cloud Services Availability. International Journal of Scientific Research and Management, 11(08), 4899. doi: 10.18535/ijsrm/v11i08.ec03

APPENDIX A: QUANTITATIVE PERFORMANCE DATA

A.1 Baseline SOC Operational Metrics (Pre-Intervention)

The following metrics represent the operational state of the Konza National Data Centre (KNDC) SOC during the six-month monitoring period prior to the CSA model implementation.

Metric	Mean ± SD	Range	Key Observation
Monthly Alert Volume	6,070 ± 541	5,420–6,780	25% growth trend
True Positive Rate	1.6% ± 0.14%	1.5–1.6%	Stable but extremely low
False Positive Rate	98.41% ± 0.052%	98.35–98.48%	Primary driver of alert fatigue
MTTD (hours)	199.67 ± 8.57	187–212	Worsening trend (+13.4%)
MTTR (hours)	376.0 ± 23.88	342–410	Worsening trend (+19.9%)
Analyst Workload Index	88.3 ± 2.5	85–92	Increasing strain (+8.2%)

A.2 Primary Outcome Measures: Comparison Table

Comparison of baseline performance versus CSA-enhanced performance across 250 controlled trials.

Outcome Measure	Baseline Mean	CSA Enhanced	Improvement	Cohen's d
MTTD (min)	192.89 ± 58.3	82.88 ± 38.9	57.03% (110 min)	5.93 (Very Large)
MTTR (min)	348.99 ± 93.3	165.19 ± 72.5	52.67% (183 min)	4.71 (Very Large)
FPR (%)	72.29 ± 15.3	19.70 ± 8.6	72.75% reduction	8.68 (Very Large)

Confidence (1-5)	2.98 ± 0.76	4.48 ± 0.59	50.32% increase	3.96 (Very Large)
Cognitive Load (TLX)	78.61 ± 18.2	45.04 ± 15.9	42.69% reduction	3.20 (Large)

APPENDIX B: STATISTICAL ANALYSIS

B.1 Two-Way ANOVA: Impact of Attack Type vs. Condition

An analysis of variance (ANOVA) was conducted to determine if the CSA model's benefits were consistent across different cyberattack categories.

Source of Variation	df	F-statistic	p-value	η2 (Effect Size)
Attack Type	4	312.46	< 0.001	0.249 (Large)
Condition (CSA)	1	490.37	< 0.001	0.490 (Large)
Interaction	4	9.27	< 0.001	0.037 (Small)

Note: The low interaction effect ($\eta^2 = 0.037$) confirms that the CSA model provides consistent performance gains regardless of the attack mechanism.

B.2 Experience Level & Performance Leveling Effect

Analysis of how the CSA model reduces the performance gap between novice and expert analysts.

Experience Level	N	Baseline MTTD	CSA MTTD	Improvement (%)
Level 1 (1-2 years)	2	195.4 min	87.2 min	55.4%
Level 2 (3-5 years)	2	191.2 min	82.5 min	56.9%
Level 3 (5+ years)	2	148.9 min	73.1 min	50.9%

Performance Variance Reduction: The variance in analyst performance dropped by **61%** (from σ 23.3% to 9.1%).

APPENDIX C: QUALITATIVE THEMATIC ANALYSIS

Key themes derived from 20 semi-structured interviews with SOC personnel to contextualize quantitative gains.

Theme 1: Alert Fatigue Mitigation: Senior analysts reported that the reduction of the FPR from 72% to 20% allowed them to move past "background noise" to focus on actionable threats.

Theme 2: Data Silo Integration: Analysts noted that the Perception layer's automated correlation saved approximately 30 minutes of manual tool-switching per incident.

Theme 3: Response Safety: The tiered response model (human-in-the-loop for high-impact actions) mitigated fears of accidental outages in high-availability environments.

Theme 4: Cognitive Scaffolding: Junior analysts cited "explainable AI" and confidence scores as tools that helped them understand senior-level reasoning patterns.

APPENDIX D: SPSS AND NVIVO ANALYSIS

D.1 SPSS Quantitative Validation Strategy

The quantitative data was processed using **SPSS (Statistical Package for the Social Sciences)** to ensure the reliability and validity of the experimental results.

Paired-Sample T-Tests: SPSS was utilized to compare the baseline (Pre-CSA) and experimental (Post-CSA) means for all 250 observations. This confirmed that the 57.03% improvement in MTTD and 52.67% improvement in MTTR were statistically significant at the $p < 0.001$ level.

Effect Size Calculation: Beyond simple significance, SPSS was used to calculate Cohen's d to determine practical significance. The analysis revealed a "Very Large" effect size for False Positive Rate reduction ($d = 8.68$), indicating a transformational impact on SOC operations.

Two-Way ANOVA: This test was performed in SPSS to analyze the interaction between the CSA Condition and Attack Type. The results ($R^2 = 0.776$) indicated that the model explains 77.6% of the variance in detection times, with a negligible interaction effect ($\eta^2 = 0.037$), proving the model's robustness across different threat categories.

Performance Leveling Analysis: SPSS descriptive statistics and variance tests (Levene's test) confirmed a 61% reduction in performance variance among analysts, statistically validating the "leveling effect" where junior analysts approached expert-level performance.

D.2 NVivo Qualitative Thematic Synthesis

The qualitative data from 20 semi-structured interviews were analyzed using NVivo to identify recurring patterns and human-centric impacts of the CSA model.

Thematic Coding: Interview transcripts were imported into NVivo and coded into five major themes: Alert Fatigue, Data Silos, High-Availability Tension, Skills Gaps, and Automation Opportunities.

Word Frequency and Nodes: NVivo queries highlighted "Cognitive Load" and "Tool Switching" as high-frequency pain points in the baseline environment. This qualitative evidence directly contextualized the 42.69% reduction in NASA TLX scores observed in the quantitative phase.

Cross-Case Analysis: NVivo allowed for the comparison of responses between Junior (Level 1) and Senior (Level 3) analysts. This analysis revealed that while all benefited, Level 1 analysts specifically credited the "Explainable AI" features for their increased confidence ($+50.32\%$), as they could finally see the "reasoning patterns" of more experienced staff.

Qualitative-Quantitative Triangulation: By using NVivo to map analyst quotes to specific KPI improvements, the research achieved triangulation. For example, the "Perception Layer" node in NVivo contained multiple references to "jumping between systems," which explains the technical driver behind the 110-minute reduction in MTTD.

D.3 Summary of Integrated Methodology

The dual-software approach ensured that the SPSS findings provided the "what" (statistically significant performance gains), while the NVivo analysis provided the "why" (reduced mental pressure, improved tool integration, and better decision support). This mixed-methods rigor supports the generalizability of the CSA model for other high-availability critical infrastructure environments