Critical Care, Critical Defense: Dissecting Hospital Security Challenges to Advance Attack Detection

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Simon Benedikt Weber

geboren in

Wuppertal

Düsseldorf, Januar 2025

aus dem Institut für Informatik
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Berichterstatter:

1. Prof. Dr. Martin Mauve
2. Prof. Dr. Michael Schöttner
3. Prof. Dr. Michael Pilgermann

Tag der mündlichen Prüfung: 10. April 2025

You can't defend. You can't prevent.

The only thing you can do is detect and respond.

– Bruce Schneier

Abstract

This dissertation addresses the critical challenges of implementing effective attack detection in hospitals, an industry increasingly targeted by sophisticated cyber attacks. Hospitals are complex, heterogeneous environments where IT systems, medical technology, and utility systems operate in interdependent ecosystems, often without sufficient observability or protection. The IT Security Act 2.0 in conjunction with federal guidelines establishes clear requirements for *state of the art* attack detection. However, these guidelines are largely generic and fail to account for the specific constraints of hospital infrastructures. This work aims to bridge this gap by aligning legal compliance, scientific advancements, and practical implementation to advance hospital security.

A comprehensive assessment of the present state of attack detection measures through a nationwide survey of German hospitals and expert interviews reveals critical gaps, particularly in industry-specific technology. Vendor restrictions, outdated systems, and limited logging capabilities hinder integration into conventional detection infrastructures. To address these shortcomings, this dissertation explores the research landscape for detecting attacks on medical cyber-physical systems by conducting a systematic literature review. The review identifies advancements in network-based anomaly detection and the integration of machine learning-based approaches while highlighting limitations such as a lack of public datasets and the immaturity of domain-specific detection methods.

Building on these findings, this dissertation introduces measures to advance attack detection in hospitals. The use of honeypots as a substitution for traditional solutions is explored. The work further investigates the potential of Large Language Models to enhance honeypot scalability, interaction realism, and adaptability to diverse systems and devices. Additionally, Manufacturer Disclosure Statements for Medical Device Security are evaluated as a structured tool to analyze the IT security landscape of medical devices. When systematically assessed, they provide a valuable reference point for security measures and future research endeavors. These proposals culminate in a comprehensive, multi-layered approach for hospital attack detection that addresses current gaps while paving the way for long-term advancements.

The results of this dissertation now serve as a basis for revising the industry-specific security standard for hospitals, ensuring that the *state of the art* specifications are not only compliant with regulatory requirements but also feasible and effective within the hospital context. By combining empirical analysis, scientific insights, and practical recommendations, this work provides a framework for advancing attack detection in hospitals. It contributes to improving the resilience of hospitals while guaranteeing operational continuity, integrity of healthcare infrastructures, and ultimately patient safety.

Acknowledgements

First of all, I would like to thank Martin Mauve, whose guidance and encouragement allowed me to explore paths beyond the standard curriculum long before this thesis began. He opened up the field of IT security during my Bachelor's studies and supported me in pursuing a specialization not envisaged in the Master's program. Thanks to the unique job arrangement he facilitated, I was able to dedicate half of my working time during my doctorate to the topics of this dissertation and the other half to enhancing the university's IT security. This arrangement created a synergy that enriched both areas. Martin has a unique way of dealing with people. His leadership is characterized by a high level of trust, an always open ear, and explicit appreciation – always seeking to bring out the best in people. This makes working with him an extraordinary experience and I am looking forward to our time ahead.

In my search for a subject-specific mentor, I was fortunate to connect with Michael Pilgermann. He provided me with the opportunity to address an issue affecting the security of millions of patients in Germany. His invaluable professional guidance and insightful discussions have significantly shaped this work. Michael opened doors to conversations with key stakeholders across Germany and he consistently encouraged me to leave the confines of the office and engage with the real world. Beyond his professional excellence, Michael's warmth, openness, and humanity have left a lasting impression. He is the kind of mentor you can always count on – whether for insightful advice or simply sharing a beer and a good conversation. Even though we took advantage of every opportunity, I'm already looking forward to the next occasion.

I would like to thank Michael Schöttner not only for taking the time to review this thesis but also for his dedication to teaching security topics at HHU and supervising theses that align with the practical needs of the ZIM. I really enjoyed this collaboration!

I am deeply grateful to the representatives of the hospital industry who shared their time, expertise, and insights. Their openness in discussing the challenges of digital hospital infrastructures provided our working group with foundational information. In particular, I wish to thank Markus Holzbrecher-Morys, who, as spokesperson for the BAK MV, enabled us to create meaningful and practical contributions through our work. He was always available to answer all our questions about the specifics of the industry. I would also like to thank the BAK working group "State of the Art" and especially Stefan Bücken. After an intense two-day meeting of the working group in spring 2022, which Stefan described afterwards as a „baptism of fire", he immediately welcomed me as a permanent guest of the UAK. His critical questions and background explanations have honed and improved our work. Also thank you for proofreading this thesis, Stefan! The cooperation with the BwZKrhs – and here in particular André Jakob and Heiko Hubrath – not only helped me understand the practical challenges of implementing IT security in hospitals but also allowed me to witness these challenges firsthand. Their technical perspective and the live insights into hospital operations made many things more

tangible for me. Having the opportunity to "test" decommissioned medical technology meant that I not only understood what a C-arm is used for in hospitals but also what it can do that the manufacturer never intended.

I would like to thank my co-authors and colleagues. First, the MedSec Group – in particular Stefan Stein and Thomas Schrader – from Brandenburg, who always provided new ideas and supported with their enthusiasm. Secondly, Marc Feger, whose machine learning and LLM expertise made our paper what it is today. Additionally, Marc always has a good story to tell. A quality that made the collaboration even more enjoyable. Thirdly, my colleagues at the university computing center, especially in AbtKS. This other component of my unique job has helped me on many occasions to experience practical problems in the implementation of IT security firsthand. I would like to particularly highlight my colleagues from the CERT: Frank Hommes and Philipp Rehs, who taught me critical skills in handling security incidents. They showed me how to remain calm and keep a clear perspective under pressure – but also when to act decisively, because the place is really on fire. Finally, I owe gratitude to my colleagues from the chair for Computer Networks and Communication Systems, because the collegiality despite the office "on the other side" is not a given. I can wholeheartedly recommend holding seminars with Jan Steimann and Markus Brenneis. Björn Ebbinghaus deserves recognition for his sharp eye on security, which has benefited not only me but the entire HHU. Kata Boland offers advice on dissertations and disputations while fostering a supportive and cohesive team dynamic. And Lisa Lorenz not only masterfully handles organizational matters, but also has the hot tips for the rental housing market.

At last, I would like to express my gratitude to my family and friends. I wish to thank Robin Weishaupt for the many thoughtful conversations, which blend professional and private topics and have been a source of inspiration since the first HHU semesters. Also thanks for proofreading (not only) this thesis! Joseph Adams, Marcel Käufler, and Tobias Uelwer, who have not only shared numerous highs and lows with me but also a passion for great coffee. Moritz Wilke for all the great times we have had together since fifth grade. Who knows where I would be without the cumulative hours of waiting, but I'm sure my life wouldn't have been nearly as worth living. Benedikt Schmeitz and the rest of the Mana crew also contributed a lot to this. Florian Jansen and Julius Kernbach, the „gladbacher jonges" who have seen the world (and shown me part of it). Tina Krüger for all our great conversations and especially for her positive worldview; I never leave our encounters without at least one new thought-provoking insight because she always has a refreshing perspective on the world! Katharina Wehr, my fellow second-degree study sufferer, contact person in all situations and just always such a great friend since school days. Bianca Faßbender for the support and all the understanding. At every milestone of this doctorate, you brainstormed with me, discussed, often comforted, or motivated me. You not only make this work better, you also make me better. Finally, to my family, because you have accompanied and supported me the longest of all. You encouraged my interest in technology from an early age, supported me in my career choice, and, because of your trust in me, I was able to persevere when I got stuck or even turned around. That is not something to be taken for granted, and that is why I will forever be deeply grateful to you. Without you, I would not be the person I am today. Thank you.

Contents

Acronyms and Abbreviations

B3S	Branchen-spezifischer Sicherheitsstandard
BAK	Branchenarbeitskreis
BSI	Bundesamt für Sicherheit in der Informationstechnik
CPE	Common Platform Enumeration
CT	Computed Tomography
DICOM	Digital Imaging and Communications in Medicine
DIDS	Distributed Intrusion Detection System
DTK	Deception Toolkit
EDR	Endpoint Detection and Response
EHR	Electronic Health Record
ENISA	European Union Agency for Cybersecurity
GDPR	General Data Protection Regulation
HIDS	Host-Based Intrusion Detection System
HIS	Hospital Information System
HL7	Health Level 7
HL7 FHIR	Fast Healthcare Interoperability Resources
ICU	Intensive Care Unit
IDS	Intrusion Detection System
IPS	Intrusion Prevention System
IT	Information Technology
kDL	Kritische Dienstleistung
KRITIS	Kritische Infrastruktur
LIS	Laboratory Information System
LLM	Large Language Model
MRI	Magnetic Resonance Imaging
MCPS	Medical Cyber-Physical Systems
MDS2	Manufacturer Disclosure Statement for Medical Device Security
MT	Medical Technology
NIDS	Network-Based Intrusion Detection System
OH SzA	Orientierungshilfe zum Einsatz von Systemen zur Angriffserkennung

OS	Operating System
OT	Operational Technology
PACS	Picture Archiving and Communication System
RIS	Radiology Information System
SBoM	Software Bill of Materials
SIEM	Security Information and Event Management
SLR	Structured Literature Review
SOC	Security Operations Center
SSH	Secure Shell
STIX	Structured Threat Information Expression
SzA	Systeme zur Angriffserkennung
TAXII	Trusted Automated Exchange of Indicator Information
UKD	University Hospital of Düsseldorf
VPN	Virtual Private Network
VT	Versorgungstechnik
XDR	Extended Detection and Response

Chapter 1

Introduction

Hospitals and healthcare systems face an unprecedented rise in cyber security threats, driven by the growing reliance on digital technologies and interconnected networks (WHO and Statista, 2024). Around 2020, incidents predominantly involved automated malware or opportunistic attackers who leveraged human errors and often lacked explicit intent to specifically target healthcare facilities (Decker et al., 2023; ENISA, 2023). A notable example, and a key motivation for this work, is the attack on the University Hospital of Düsseldorf (UKD) in Germany in 2020. Initially, this incident was considered the first case in which a cyber attack directly caused a human fatality (Miliard, 2020). Subsequent legal proceedings, however, concluded that it could not be definitively determined that the patient's demise was attributable to the delay caused by the attack (Ralston, 2020). During incident response, it became evident that the attackers had mistakenly encrypted the hospital's systems instead of those of an adjacent university. When informed of their mistake, the attackers provided the decryption key without demanding a ransom. Nevertheless, the attack severely disrupted hospital operations and required a lengthy recovery process (Miliard, 2020). This incident exemplifies how even large, well-resourced hospitals can be seriously impacted by more or less coincidental and undirected security incidents.

Since 2022, the global threat landscape has evolved and the frequency and sophistication of attacks has further increased (Decker et al., 2023). Check Point Research (2023) reports that the healthcare sector is now among the three most attacked sectors. The European Union Agency for Cybersecurity (ENISA) specifies that in the sector particularly hospitals are affected (ENISA, 2023). In contrast to the incidental, malware-driven attacks described above, attacks from adversarial groups specializing in healthcare infiltration are now commonplace, reflecting a shift to deliberate campaigns. Targeted, multi-staged, so-called 'hands-on-keyboard' attacks, in which an adversary actively interacts with the hospital systems, are on the rise (71% of attacks were non-malware activity and interactive intrusions increased 50% in 2022) (CrowdStrike, 2023; Decker et al., 2023). Ransomware remains the predominant threat, driven by the prospect of financial gain (ENISA, 2023). Attackers' interest has increased to the point that the Health Sector Cybersecurity Coordination Center (2024) reports the top ten ransomware groups actively targeting the US health and public health sector. Due to the life-critical services hospitals provide, they are under immense pressure to restore functionality rapidly, as

operational downtime can directly threaten patient safety. This urgency makes them more willing to pay ransoms in order to accelerate recovery (Sophos, 2024).

However, financial gain is no longer the only driving force behind these attacks. A growing number of attackers are motivated by other objectives, such as undermining public trust in governmental institutions. Hospitals represent pillars of community stability, hence sustained or high-profile attacks can cause fear, uncertainty, and long-lasting damage to their reputation (Decker et al., 2023). Breaches jeopardize not just the confidentiality and integrity of patient data but also pose a long-term risk, as inaccurate records lead to diagnostic or treatment errors. High-profile incidents have demonstrated that even a short-term disruption – such as encrypting patient records, disabling critical medical equipment, or forcing staff to resort to manual processes – can impact patient care delivery (ENISA, 2023; Ponemon, 2023). A comprehensive and resilient security framework is essential to protect critical infrastructures from emerging cyber security threats. Advancing attack detection in hospitals could preserve their operational integrity and ensure patient safety.

1.1 Motivation & Problem Statement

In response to the escalating cyber threats, the German legislator revised the IT Security Act in 2021, mandating the implementation of *state of the art* attack detection systems across critical infrastructure (German: Kritische Infrastruktur (KRITIS)) organizations (Bundestag, 2021). According to the regulations, various sectors – including energy, water, food, and healthcare – are classified as KRITIS. Within the healthcare sector, hospitals are subject to KRITIS regulations if they exceed a threshold of 30,000 full inpatient cases per year (Justiz, 2016). Such institutions must demonstrate compliance with security standards. Smaller hospitals, that do not meet the KRITIS threshold, are not directly affected by the IT Security Act 2.0. However, they may still be impacted by additional legislative requirements or industry-specific regulations, as discussed in section 2.2. This extension of obligations underscores the growing acknowledgment that healthcare providers – regardless of their size – play a vital role in societal resilience.

While various standards and best practices define what constitutes *state of the art* in attack detection, these guidelines are typically generic and fail to account for the distinctive challenges posed by hospital environments. The digital ecosystems in hospitals integrate medical, administrative, and operational components, where patient safety and continuity of care add an additional layer of complexity. There is a pressing need for targeted guidance that bridges the gap between formal compliance requirements and the practical realities of hospital operations. Such industry-specific orientation is crucial to achieving meaningful and sustainable improvements in the security posture of hospitals.

1.2 Contributions

This dissertation includes four contributions to advancing hospital attack detection. First, it systematically explores the legal and regulatory frameworks governing hospital security in Germany, aligning them with industry-specific needs identified through a systematic analysis of the present state of hospital security. A structured literature review is used to pinpoint the *state of the art* attack detection for medical cyber-physical systems, establishing a reference for future research and operational improvements.

Second, the work introduces honeypots as a central element in circumventing vendor restrictions and outdated systems. By removing the dependency on native device functionalities and lengthy certification adjustments, honeypots allow for cost-effective, dynamic attack detection. Moreover, this dissertation explores the integration of Large Language Models (LLMs) as honeypot backends to automate and streamline their development. A core contribution here is the introduction of a unified evaluation framework for LLM-based honeypots, allowing researchers to conduct reproducible assessments and compare results across diverse configurations. This framework can guide future advancements in theory and practice.

Third, to address the existing knowledge gap on the security characteristics of medical devices, this dissertation offers the first systematic evaluation approach for Manufacturer Disclosure Statement for Medical Device Security (MDS2). By enabling researchers and hospital operators to derive meaningful insights from these documents, it fosters a more nuanced understanding of medical device security. Researchers can validate their assumptions against real-world data, while operators can make informed decisions on prioritizing security measures. This approach thus strengthens the link between research outcomes and operational needs, ensuring that proposed solutions resonate with actual hospital environments.

Fourth, the dissertation translates these insights into concrete recommendations, culminating in 44 evidence-based conclusions containing short-term measures and mid-term strategies. These actionable outcomes enable immediate enhancements to current security practices while guiding longer-term developments, thereby ensuring that the findings remain both practically applicable and forward-looking.

The results from this work are actively shaping the ongoing revision of the industry-specific security standard for German hospitals. By aligning regulatory requirements, cutting-edge research, and practical implementation, this dissertation lays the groundwork for more robust, evidence-based, and context-aware attack detection that maintains operational continuity and ultimately enhances patient safety.

1.3 Outline of this Thesis

Chapter 2 establishes the foundational concepts of attack detection, including the evolution of relevant technologies, and sets the legal and regulatory context. In addition, the heterogeneity of digital hospital ecosystems is examined.

Building on these fundamentals, Chapter 3 provides the conceptual frame linking the subsequent chapters. It integrates insights from the assessment of the present state and the research landscape, illustrating how to specify an industry-specific *state of the art* that reflects technical feasibility and practical needs. This chapter also discusses how honeypots and the systematic analysis of MDS2 documents can serve as short- and mid-term measures to alleviate immediate challenges until more comprehensive solutions become viable.

Next, Chapter 4 dives deeper into the scientific landscape by presenting the Structured Literature Review (SLR) focused on attack detection for Medical Cyber-Physical Systems (MCPS). This review presents key advancements, clarifies existing research gaps, and identifies pressing limitations – such as scarce datasets and the immaturity of domain-specific detection approaches – that must be addressed to advance the research area.

Chapter 5 explores the feasibility of using LLMs to reduce the complexity, development costs, and risks associated with traditional honeypots. The introduction of a unified evaluation framework enables standardized comparisons and reproducible research, thereby accelerating the development of more dynamic and effective detection mechanisms.

Subsequently, Chapter 6 presents a methodology for analyzing MDS2 documents. This approach translates the documents into actionable security knowledge, empowering researchers and hospital operators to make evidence-based decisions and strategically prioritize security measures.

Chapter 7 applies the insights, methods, and tools developed throughout the previous chapters to support the revision of the industry-specific security standard for German hospitals. By aligning these findings with regulatory requirements, this chapter demonstrates the practical value and relevance of the proposed solutions.

Finally, Chapter 8 summarizes the findings of this dissertation and reflects on how the proposed solutions address identified gaps in hospital attack detection. It also outlines directions for future research ensuring that this work lays a sustainable foundation for ongoing advancements in hospital security.

Chapter 2

Foundations and Background

The growing complexity of hospital infrastructures and the increasing sophistication of attacks necessitates a thorough understanding of the relevant technologies, frameworks, and systems for attack detection. This chapter establishes the foundation by examining attack detection technologies, the regulatory landscape governing hospital security in Germany, and the unique characteristics of digital infrastructures of hospitals.

2.1 Evolution of Attack Detection Technologies

The detection of attacks has undergone a major transformation, evolving alongside advancements in technology and the shifting threat landscape. Subsequently, the historical development of attack detection technologies is traced out, starting with early Intrusion Detection System (IDS) approaches and progressing to modern, integrated solutions. Understanding this evolution highlights the strengths and limitations of existing approaches and sets the stage for identifying their applicability to the unique challenges faced by hospital environments.

2.1.1 Early Technologies and Intrusion Detection Systems

Early development of attack detection systems began in the 1980s, driven by the growing need to secure expanding computer networks. As these networks evolved, the threat of unauthorized access and malicious activities increased, prompting research into methods for identifying and mitigating them. One of the most influential contributions in this domain was made by Anderson (1980), who introduced the concept of "audit trails" as a method to detect security violations, particularly unauthorized access attempts. Anderson's work laid the foundation for future IDS by emphasizing the importance of systematic tracking and analyzing of system activities. Signature-based IDS were among the first practical implementations of intrusion detection technology. These systems rely on predefined patterns, later known as *signatures*, which correspond to known attack vectors. One of the earliest and most influential contributions to this field was the model by Denning (1987), which proposed to use signatures to detect malicious activities within a network or system. This concept was later implemented in the Intrusion Detection Expert System (IDES), developed by Lunt et al. (1992) and one of

the first practical applications of signature-based detection. The basic operating mechanism of signature-based IDS involves scanning incoming network traffic or monitoring system processes to identify data patterns that match known signatures of malware or unauthorized access attempts. When such a pattern is detected, the IDS generates an alert, allowing administrators to respond to the potential threat. While effective at identifying known threats, these systems have a critical limitation: they can only detect attacks that have already been observed and cataloged. This reliance on existing signatures made them vulnerable to attacks, where new or previously unknown exploits would bypass the detection system until signatures were updated. With the use of IDS the motivation of attackers increased to introduce small changes to existing malware. The minor modifications prevent the recognition as known malware, making it appear to signature-based IDS as benign activity. An ongoing race between attackers and defenders emerged, with attackers constantly tweaking their malware to outsmart detection, and defenders continually updating their detection systems to keep pace. Since the effectiveness of signature-based systems is limited by their inherent lag in identifying such nuanced changes, a rapid and resource-intensive update cycle is needed. This requirement for constant signature updates placed a maintenance burden on organizations and their security teams (Khraisat et al., 2019).

Anomaly-based IDS emerged as a response to the limitations of signature-based systems. Instead of looking for known attack signatures, anomaly-based IDS focus on detecting deviations from established baseline behavior. These systems create a model of normal activity for a network or system, encompassing metrics such as typical network traffic patterns, user behaviors, and system process behavior. Any significant deviation from this baseline is flagged as a potential intrusion, as it may indicate unauthorized or malicious activity. The main advantage of anomaly-based IDS lies in their ability to detect unknown, so-called 'zero-day' threats, which may not follow conventional attack patterns. Additionally, these systems are less dependent on regular updates compared to signature-based IDS. However, anomaly detection is accompanied by its own set of challenges, particularly the issue of high false-positive rates. Since benign activities can sometimes deviate from established baselines, the IDS may incorrectly flag legitimate actions as intrusions, leading to alert fatigue among administrators. All systems and networks have unique behavior patterns which may change over time or during specific events. For instance, the publication of annual financial statements is a one-time yearly event generating numerous deviations from typical behavior. These temporary deviations can cause anomaly-based systems to misinterpret legitimate activity as a threat, especially in complex, heterogeneous environments. The difficulty of accurately distinguishing between genuine attacks and benign anomalies challenges the efficiency of such systems (Khraisat et al., 2019).

Both, signature-based and anomaly-based IDS can target different layers of the IT infrastructure: the network or a specific host. These are referred to as Network-Based Intrusion Detection System (NIDS) and Host-Based Intrusion Detection System (HIDS) (Mukherjee et al., 1994). An NIDS monitors network traffic in real-time, analyzing data packets across entire network segments. Positioned at strategic points, for example between internal networks and external internet gateways, an NIDS can detect suspicious activity across a broad array of devices. Their ability to provide visibility makes them particularly beneficial in heterogeneous

networks, including those containing devices with limited native security features. An HIDS operates on the individual device level, focusing on monitoring the internal behavior of a specific host – such as a server, or a workstation. By analyzing logs, file integrity, and system processes, an HIDS can detect anomalies or malicious activities confined to a particular device. They can be effective for protecting critical devices that store sensitive data or for safeguarding high-value equipment connected to a network. Typically, an HIDS must be installed directly on the system it monitors. That can be challenging for various reasons, such as compatibility issues, resource constraints, or concerns regarding system performance impacts. However, a significant benefit of an HIDS compared to an NIDS is that it often provides deeper visibility into system activities. On a host, one can observe far more detailed information compared to the network layer, such as specific processes, file integrity, and user actions, even when network traffic is encrypted (Knerler et al., 2023).

The Distributed Intrusion Detection System (DIDS), introduced by Snapp et al. (1991), emphasizes the importance of integrating data from multiple sources for detecting and responding to threats. DIDS utilize distributed agents and a centralized analysis component to correlate data across networks, creating an early framework of collaborative threat detection.

An Intrusion Prevention System (IPS) is designed to not only detect but actively prevent unauthorized access, misuse, or attacks within a network. Unlike IDS, which only monitor and alert administrators about potential security breaches, IPS go a step further. By taking automated actions to block or mitigate detected threats in real time, they combine detection capabilities with proactive enforcement. IPS typically operate inline with network traffic, allowing them to analyze, identify, and react to malicious activities by dropping packets, blocking IP addresses, or reconfiguring network policies (Fuchsberger, 2005).

In general, organizations require more advanced and integrated solutions than traditional IDS. Recent solutions combine the strengths of signature and anomaly detection methods, integrating data from multiple sources to provide a more comprehensive approach to threat detection and prevention. The advancement of IDS and IPS laid the groundwork for the development of such solutions.

2.1.2 Recent Advancements in Attack Detection

Security Information and Event Management (SIEM), Endpoint Detection and Response (EDR), and Extended Detection and Response (XDR) represent a shift toward a more unified, real-time approach for detecting and responding to security incidents. In this context, the term *endpoints* is increasingly used instead of hosts to reflect a broader and more modern perspective on IT security architectures. In addition to traditional servers and workstations, this also includes mobile devices, Operational Technology (OT), and cloud-based resources. The term endpoint thus highlights the versatility and diversity of today's digital environments.

SIEM emerged as a response to the demand for more thorough security monitoring compared to isolated IDS. While IDS were capable of detecting individual incidents, they lacked the abil-

ity to correlate data across multiple systems to offer a unified view of an organization's security posture. SIEM systems fill this gap by aggregating log and event data from various infrastructure components, such as firewalls, servers, databases, and network devices. They allow for cross-system comparisons, and further enrichment through correlation and analysis (Bhatt et al., 2014). This enables security analysts to detect complex attack patterns that might involve multiple, seemingly unrelated events. However, SIEM systems can create considerable overhead. Each system that should submit logs to the SIEM must be added separately, and specific information and logs to be forwarded must be identified individually for every system. Additionally, submitted logs often need to be normalized to ensure consistency. Therefore, SIEM systems still require careful tuning and regular optimization to avoid false-positive alerts and enable effective management of the large volume of incoming data.

EDR systems were developed to address the limitations of traditional detection mechanisms in securing individual endpoints. They integrate deeply into the operating system and use kernel callbacks and event tracing, to provide an EDR system with comprehensive visibility into system activities and enable real-time threat detection. The need for such low-level access to system resources and events requires a persistent presence on the endpoint, mainly in the form of installed agents or software components. Recent EDR systems use a combination of signature-based and anomaly-based detection methods, allowing them to identify both known threats and novel attacks that deviate from normal behavior. The response component signifies that upon detecting a threat, EDR can isolate compromised endpoints (like an IPS), terminate malicious processes, or even roll back systems to a previously known safe state (Karantzas and Patsakis, 2021). This capability for immediate containment of threats can prevent the spread of malware or ransomware across IT systems.

XDR build on the foundations of SIEM, EDR, and DIDS. They provide a more holistic and integrated approach to threat detection by correlating data across multiple security layers, including endpoints, networks, cloud environments, and even user behavior. XDR systems aim to eliminate the silos between different security tools, enabling cross-layered detection and response (George et al., 2021). For example, by utilizing XDR, a security analyst could not only identify an initial endpoint compromise (as with an EDR) but also track lateral movement across the network (as with a SIEM), gaining a comprehensive view of the attack. While conceptually similar to DIDS, the earlier version primarily focused on network-level integration and was constrained by the technological limitations of its time. It lacked the broader, multi-environment perspective now promised by XDR. Yet, it is important to note that XDR is a marketing term, as similar functionality can be achieved by integrating all relevant sources – such as EDR, network sensors, firewalls, and others – into a central SIEM, ensuring that all data is normalized and can be comprehensively searched. However, the promise of XDR encompasses an easier and more straightforward implementation as well as a faster state of readiness for organizations compared to a custom-built SIEM solution. Summarizing, XDR and EDR appear particularly well-suited for environments with a homogeneous digital infrastructure, where consistent endpoints and network components simplify integration and visibility. Broad endpoint coverage enhances their effectiveness by ensuring that activities across all devices

can be monitored comprehensively. As with EDR, XDR platforms often include automated response capabilities, such as isolating endpoints upon detection of malicious activities.

Threat intelligence involves gathering and analyzing data about potential threats from external sources, which helps to identify new vulnerabilities and attack vectors. Standards like Structured Threat Information Expression (STIX) and Trusted Automated Exchange of Indicator Information (TAXII) are used to format and share such intelligence across different systems (Wagner et al., 2019). By incorporating threat intelligence feeds into SIEM or XDR systems, the latest information on threat actors, tactics, techniques, and procedures can be utilized.

Security Operations Centers (SOCs) are centralized units where security analysts monitor, detect, and respond to security incidents across an organization's entire digital infrastructure (Knerler et al., 2023). SOCs utilize all the aforementioned tools, processes, and intelligence, enabling real-time monitoring and efficient response to potential threats. Due to the overhead inherent to operating the systems, including the integration of multiple systems into a SIEM, normalization of logs, and careful tuning of alerts, SOCs are necessary to manage and make sense of the large volumes of data being generated. By providing a dedicated team of experts and a structured approach to handling incidents, SOCs ensure that threats are identified promptly and appropriate measures are taken immediately to mitigate risks.

2.2 Regulatory and Legal Landscape for IT Security in German Hospitals

The regulatory landscape governing IT security in German hospitals is shaped by increasingly stringent legal requirements. The IT Security Act 2.0 (Bundestag, 2021) mandates the implementation of security measures to safeguard KRITIS organizations, including attack detection systems (German: Systeme zur Angriffserkennung (SzA)). The German legislator introduced the term SzA to provide a more precise and comprehensive designation for systems that detect various forms of cyber attacks. Unlike the traditional term IDS, which linguistically focuses on unauthorized access or intrusions, SzA encompasses the detection of a broader range of attack activities, including malware, abuse of legitimate access, and other anomalous behavior. This terminology recognizes the evolving nature of threats and improves clarity in legal and regulatory frameworks. Subsequently, the legal framework is outlined, focusing on the key changes for SzA.

2.2.1 The IT Security Act 2.0 and SzA

The IT Security Act 2.0, enacted in May 2021, is an extensive update to the original legislation from 2015. As part of Germany's broader efforts to secure vital services, KRITIS refers to infrastructure whose disruption would have a profound impact on societal functions. Healthcare, as a critical sector, is subject to heightened scrutiny under this framework. The

act introduces stricter requirements for adopting organizational and technical measures that meet the *state of the art*, aiming to ensure the availability, integrity, and confidentiality of sensitive data and systems. One of the most impactful updates introduced by the IT Security Act 2.0 regarding the protection of KRITIS is the requirement for SzA. It is important to state that a legislative proposal to implement the EU NIS-2 directive is currently negotiated in the Bundestag (Bundestag, Deutscher, 2024). This directive relocates the explicit obligation to use SzA, yet only the phrasing is slightly adjusted. The substance of the regulatory requirement and, hence, the SzA intention remains virtually unchanged. The act imposes mandatory reporting of significant IT security incidents to the regulatory authority, ensuring real-time monitoring and response to threats across the critical infrastructure sectors. In Germany, the regulatory authority responsible is the Federal Office for Information Security (German: Bundesamt für Sicherheit in der Informationstechnik (BSI)). While SzA is not legally detailed, the BSI has issued a guidance document (German: Orientierungshilfe zum Einsatz von Systemen zur Angriffserkennung (OH SzA)) for KRITIS operators (Bundesamt für Sicherheit in der Informationstechnik, 2022).[1]

The OH SzA outlines three essential components that are needed to fulfill the requirements for attack detection: logging, detection, and response. Briefly summarized, SzA must continuously collect (logging) and evaluate information to identify security-relevant events (detection). To prevent disruptions resulting from attacks SzA must also be able to respond to these attacks effectively (response). The OH SzA further specifies, that all systems, components, or processes that are essential for delivering the "critical service" (German: Kritische Dienstleistung (kDL)) must be covered. The kDL involves every service provided by a KRITIS organization to the general public, where failure or impairment could lead to significant bottlenecks or pose threats to public safety (Justiz, 2016). The passage of the OH SzA, thereby, explicitly includes Information Technology (IT) and OT systems, as well as embedded systems. The OH SzA does not mandate specific technologies to be used in fulfilling these requirements. Instead, it describes the functional objectives that must be achieved. However, the descriptions are sufficiently detailed and contextually aligned to suggest, with a high degree of certainty, that the technologies mentioned in the previous section – namely EDR, XDR, and SIEM – are intended. These technologies provide the necessary capabilities for logging, detecting, and responding to security incidents. In cases where EDR/XDR solutions cannot be implemented, the OH SzA implicitly acknowledges alternative fallback options. For instance, leveraging existing log-forwarding capabilities to integrate endpoints into a central SIEM ensures logging and detection capabilities. A following fallback is network-level monitoring, such as deploying NIDS, which log network traffic and detect threats through analysis of this traffic. For endpoints where all other measures are infeasible, unspecified alternatives may be adopted as a final fallback, provided they achieve a comparable security posture and are sufficiently justified.

The OH SzA is designed for all KRITIS sectors and organizations. It is detailed, yet generic enough to be reasonably suitable for all sectors. If a sector has specific requirements or chal-

[1] In July 2022, the MedSec-Group provided comments on the community draft of the OH SzA, which can be accessed online: https://www.th-brandenburg.de/fileadmin/user_upload/fb-informatik/bilder/Security_Forensics/StellungnahmeTHB.pdf

lenges, the legislation allows the industry to formulate its own *state of the art* specifications based on its expertise via a so-called Industry-Specific Security Standard (German: Branchenspezifischer Sicherheitsstandard (B3S)) (Bundesamt für Sicherheit in der Informationstechnik, 2024a). Therefore, the B3S represents a critical framework for enhancing IT security in KRITIS sectors, tailored specifically to meet the unique needs of individual industries. B3S are developed and maintained by working groups formed by KRITIS operators, security experts, and regulatory agencies. These are organized through the public-private partnership UP KRITIS (Bundesamt für Sicherheit in der Informationstechnik, 2024b). The committee structure within the UP KRITIS provides for the formation of industry working groups (German: Branchenarbeitskreis (BAK)) to address sector-specific issues and develop tailored concepts (Bundesamt für Sicherheit in der Informationstechnik, 2024c). B3S eligibility is determined by the BSI for each version and at fixed intervals. This way, legislation has ensured that the B3S is dynamic enough to ensure that KRITIS organizations maintain compliance with the latest security practices (Bundestag, 2009) while leaving enough flexibility to account for industry-specific challenges.

2.2.2 *B3S Hospital*: The Industry-Specific Security Standard for Hospitals

The B3S for hospitals was initially introduced following the IT Security Act 1.0 in 2015, but its scope has since expanded to incorporate the stricter requirements of the IT Security Act 2.0. The most recent version 1.2, released in December 2022, reflects these updates (Deutsche Krankenhausgesellschaft, 2022), including the incorporation of SzA requirements. Developed by the BAK for hospitals, great care was taken, that the B3S ensure its application across diverse hospital environments, ranging from small regional facilities to large university medical centers. KRITIS hospitals that fully implement the B3S for hospitals evidently and with legal certainty meet the requirements of the IT Security Act 2.0 (Bundesamt für Sicherheit in der Informationstechnik, 2024a). Achieving certification provides protection under the act, as it demonstrates that hospitals have taken appropriate measures to secure their digital infrastructure. Yet, the implementation of the B3S is not mandatory. KRITIS Operators can choose to implement it, but they can also meet the legal requirements through other means.

For non-KRITIS hospitals, the legal framework differs, as they are not mandated to comply with the requirements of the IT Security Act 2.0. However, the enactment of the *Protection of Electronic Patient Data in the Telematics Infrastructure Act* in 2020 brought a substantial change. Through an amendment to Social Code V (Section 75c SGB V), hospitals are now required to implement *state of the art* organizational and technical measures to ensure the availability, integrity, and confidentiality of their systems. The law explicitly references the B3S as a recommended standard for fulfilling these requirements (Bundestag, 2020). In March 2024, the *Act to Accelerate the Digitalization of the Healthcare System* came into force, relocating regulations for IT security in hospitals to Section 391 of SGB V, effectively replacing Section 75c. Paragraph 1 of Section 391 reiterates the requirement to adhere to the *state of the art* for IT security. Additionally, paragraph 4 specifies that compliance with the B3S is recognized as fulfilling these obligations (Bundestag, 2024). Unlike the IT Security Act for KRITIS, the

provisions of SGB V do not explicitly mandate the implementation of SzA. Consequently, the extent to which non-KRITIS hospitals will be required to implement SzA will depend on the design and evolution of the B3S. It is plausible that a (potentially less stringent) adaptation of SzA could become part of the *state of the art* requirements for smaller hospitals in the near future. While this dual role of the B3S poses challenges for its conception and updating process, it is also one of the key strengths of the B3S, allowing hospitals to tailor the guidelines to their specific operating environment. For example, small regional hospitals with limited resources may prioritize basic security measures such as network segmentation, while larger university hospitals may implement advanced solutions. This adaptability ensures that hospitals of varying sizes and complexities can comply with legal requirements, based on risk assessments and resource availability, without compromising their operational effectiveness.

The current version 1.2 of the B3S for hospitals is implemented by most hospitals in Germany, and its specificity and clarity are praised by many (Weber et al., 2024a). Overall, the B3S is well suited for adapting regulatory requirements to the specific needs of the healthcare industry and excels in many areas of security. However, in the absence of nationally or internationally applicable industry-specific good practices with regard to SzA, the BAK has strongly leaned on the OH SzA and, thus, to cross-sector, general recommendations in the current B3S version. Currently, hospitals face the problem that these attack detection measures are only partially applicable to the specific circumstances in German hospitals. Therefore, the implementation of SzA poses a challenge for hospitals, as discussed in the next chapter. For this reason, the BAK has planned a fundamental revision of the B3S to account for the industry's specific challenges related to SzA.

2.3 The Digital Infrastructure of Hospitals

The B3S describes the complex and heterogeneous digital landscape of hospitals with its components and systems relevant to the kDL in great detail (Deutsche Krankenhausgesellschaft, 2022). It also addresses the industry-specific risk situation, explains dependencies, and measures to secure systems. The three domains relevant to SzA, based on the current B3S, are: IT, Medical Technology (MT), and utility systems (German: Versorgungstechnik (VT)).[2] Devices are becoming increasingly interconnected within and across these domains in hospitals. While the integration of these systems offers numerous operational advantages, it also introduces new security vulnerabilities that must be managed carefully to ensure the safety of patients and hospital staff. These domains are integral to the kDL and differ considerably from each other as well as from the technology and conditions of other industries.

The IT of hospitals encompasses systems that manage administrative tasks, clinical workflows, and data storage. Core components of IT systems in hospitals can be divided into two categories: general administrative IT and specialized hospital IT systems. General administrative

[2] In contrast to the B3S, this thesis does not specifically address telecommunication technology. While the distinction is reasonable in other areas of security, it is not necessary for the exploration of SzA, as these systems are not fundamentally different in other industries.

IT includes tools such as PCs and servers that support administrative tasks and communication. More recently, one can also observe an increasing reliance on cloud services in the IT domain (Weber et al., 2024a). Examples of specialized hospital IT systems include the Hospital Information System (HIS), the Radiology Information System (RIS), the Laboratory Information System (LIS), and the Picture Archiving and Communication System (PACS). The HIS serves as the central platform for managing patient records, clinical workflows, and hospital operations, while RIS and LIS manage radiological and laboratory data respectively, facilitating efficient diagnostics and results sharing. The PACS provides centralized storage and retrieval for medical imaging, enhancing access for diagnostics. These specialized hospital IT systems are responsible for storing and managing sensitive patient data, facilitating communication between departments, and supporting critical services such as scheduling, billing, and resource allocation. Furthermore, IT systems play a central role in the integration of other domains, such as medical devices and operational technologies, ensuring seamless communication and data exchange across a hospital's infrastructure (Deutsche Krankenhausgesellschaft, 2022).

MT in hospitals incorporate a wide range of medical devices critical to diagnosing, monitoring, and treating patients. It includes devices such as ventilators, infusion pumps, patient monitoring systems, and imaging systems (e.g., Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scanners). Modern medical devices are increasingly network-connected, allowing integration with hospital IT systems to facilitate data exchange, diagnostic information sharing, and real-time patient monitoring. To emphasize the network capability – and necessity – even more clearly, these systems are increasingly referred to as Medical Cyber-Physical Systems in the literature (Weber et al., 2023). In contrast to administrative IT systems, MCPS are deeply embedded in clinical operations, supporting direct patient care, and can involve physical interaction with patients. Due to the diversity of manufacturers and standards, MCPS often operate with a wide range of protocols and compatibility requirements, which creates challenges for maintenance and interoperability. Medical devices are designed for durability and consistent clinical performance, which requires high-quality materials and manufacturing standards. The focus on reliability drives up costs, making frequent replacements financially impractical. Therefore, they can have a long lifetime ranging between ten and 20 years (Seo et al., 2022). Additionally, medical devices must meet strict regulatory requirements, which include undergoing certification processes. This means that any updates or modifications often require manufacturer intervention, either remotely or with an on-site technician to perform maintenance tasks. As a result, medical devices are black boxes for network operators, who typically do not have administrative rights or even knowledge of what software is running on the devices. Some medical devices require a constant supply of materials or consumables to operate correctly, e.g., specific gases or fluids. Monitoring these supplies is crucial, and in many cases, this task is managed remotely by the manufacturer as part of a service agreement. Maintenance services and operational monitoring are the reasons why manufacturers of medical devices frequently have specific requirements for remote connections to hospital networks. These connections are typically handled through Virtual Private Network (VPN) connections (Deutsche Krankenhausgesellschaft, 2022). However, such requirements are fre-

quently neglected during purchasing phases, either because physicians prioritize the clinical functionality of a device or due to the limited availability of devices meeting specific clinical needs. Conclusively, devices are purchased without consideration for the type of service connection required, leaving network operators responsible for finding secure methods to manage such connections with external networks.

VT manages the physical infrastructure of hospitals, including systems for building management, energy supply, heating, ventilation, and air conditioning (HVAC), water supply, and medical gas systems. VT systems are crucial to ensure that hospitals remain functional by providing the essential utilities and support services that sustain the physical environment required for patient care. They maintain the appropriate conditions to operate sensitive medical equipment, which includes climate control, supply systems for medical gases, and infrastructure for essential services. Increasingly, VT systems are integrated with hospital IT infrastructure for remote monitoring and control, enabling hospitals to optimize energy consumption and automate operational efficiency. For example, HVAC systems are connected to IT networks to maintain optimal temperature and humidity levels in critical areas such as operating rooms. Similarly, energy management systems guarantee continuous power supply to essential medical devices in the event of power outages. Systems such as digital access control, fire alarms, and elevator services also fall under VT, ensuring both safety and convenience for hospital operations. VT systems, while often less recognized compared to IT and MT, are key utility systems that support the entire healthcare infrastructure (Deutsche Krankenhausgesellschaft, 2022). And even if it is not as obvious for VT as for MCPS, exploited vulnerabilities in VT could still disturb essential hospital services. For instance, disrupting elevator services might bring parts of a hospital to a halt, as elevators are crucial for transporting patients who cannot leave their beds, such as those being moved to or from operating rooms. Therefore, many VT systems are integral to the kDL. A wide range of operational restrictions mentioned in connection with MT also apply to VT. For example, VT maintenance tasks also frequently require the intervention of the manufacturer. Likewise, remote maintenance channels or regular on-site visits by engineers are common.

The heterogeneity of systems in the IT, MT, and VT domains in hospitals results in a high complexity of securing these environments. Their interconnectedness further amplifies this complexity and leads to an interdependent digital ecosystem. As a result, ensuring effective security in hospitals requires a holistic approach.

Chapter 3

Advancing Attack Detection in Hospitals

By May 2023, all KRITIS organizations, including hospitals classified as KRITIS, were legally required to have fully operational SzA systems. Non-compliance with the IT Security Act 2.0 is met with substantial penalties, including fines up to 20 million euros or 4% of global annual turnover (Bundestag, 2021). Yet the fundamental differences between the digital infrastructure of hospitals and other sectors (in terms of IT, MT, and VT, as discussed in Chapter 2) aggravate the adoption of detection measures from generic guidelines like the OH SzA. The current version of the B3S is not yet able to provide a remedy either, as it is strongly based on the OH SzA. Therefore, the upcoming version of the B3S must further specify the *state of the art* in attack detection for the hospital industry, despite the ongoing absence of industry-specific standards and good practices to build upon.

3.1 Inferring the State of the Art in Detecting Attacks

When searching for suitable ways to further specify the *state of the art*, it is instructive to refer to the Handbook of Legal Formality (German: Handbuch der Rechtsförmlichkeit, Bundesministerium der Justiz (2024)). This handbook provides a framework to infer the *state of the art*. According to this, a distinction is made between three technical standards using general clauses. These levels define the progression of technical measures from established best practices to cutting-edge innovations. The *state of the art* ranges between the *generally accepted rules of technology* (German: Allgemein anerkannte Regeln der Technik) and the *state of science and technology* (German: Stand der Wissenschaft und Technik).

The *generally accepted rules of technology* represent common practices that are widely established and have proven effective over time. Whereas, the *state of science and technology* entails the cutting edge of technological innovation, encompassing approaches that are currently discussed in research or are in an early developmental stage. The assessment of the present state of SzA implementation in German hospitals and the *state of research and technology*, allows to approximate the *state of the art* for SzA in hospitals. The underlying assumption is that the present state of SzA implementation ranges between the *generally accepted rules of technology* and the *state of the art*. Determining the *state of science and technology* for attack detection

aids in estimating the applicability of recommended measures and in identifying alternative and more feasible methods.

3.1.1 Assessing the Present State: SzA in Hospitals

To assess the present state of SzA within German hospitals this chapter is based on the results of a survey as well as interviews with industry experts (Chapter 7). Representatives of 132 hospital sites participated in the survey. The respondents included large urban medical centers and smaller rural hospitals, providing an elaborated view of the present state of SzA implementation across different types of healthcare institutions. The interviewees were industry experts and security professionals with the ability to discuss the industry challenges in detail.

Despite the regulatory mandate, the implementation of SzA in German hospitals is inconsistent, with clear differences between technological domains. Benefitting from mature, applicable frameworks and better integration capabilities, IT systems are generally more advanced (but not yet completed) in adopting SzA. In contrast, MT systems face challenges and VT systems lag further behind. Even hospital operators that self-assessed to meet OH SzA attack detection requirements in MT and VT do not yet have fully functional attack detection in these domains. This discrepancy clearly indicates the need for industry-specific adaptations, as the cross-industry recommendations are not only infeasible for most operators but also do not provide sufficient security when implemented. The following section describes challenges that emerged during the assessment of the present state of SzA in German hospitals.

3.1.2 Challenges in Securing Hospital Systems

To begin, it is essential to address the general challenges that distinguish hospitals from other industries, particularly in their security posture and exposure situation. Hospitals are open environments where untrusted individuals, such as patients or visitors, can gain physical proximity to critical systems, increasing the risk of tampering or unauthorized access. These *insider threats* add complexity to the security landscape, distinguishing the hospital industry even more from other critical infrastructures (U.S. Department of Health and Human Services, 2023). Additionally, hospital IT departments often operate with comparatively more limited budgets, restricting their ability to invest in the latest security technologies or to hire specialized security personnel. This scarcity of resources can lead to gaps in security monitoring, delayed patching, and inadequate incident response capabilities. The handling of data privacy and regulatory compliance further complicates the security landscape. Hospitals process large volumes of sensitive patient information, and maintaining compliance with regulations such as the General Data Protection Regulation (GDPR) is a continuous challenge. Ensuring that patient data is securely stored, processed, and transmitted requires implementing sophisticated security measures. These measures must not disrupt the workflow of healthcare providers, who need quick data access to deliver timely care. Furthermore, network traffic and log data may

contain sensitive patient data, too, and therefore, security teams need to anticipate this during the design and operation of SzA.

In addition to these general difficulties, the heterogeneous nature of hospital systems requires the integration of attack detection tools and the management of challenges specific to medical and operational technologies, to protect hospital infrastructure from evolving threats. Each domain IT, MT, and VT has unique security requirements and protocols, complicating the development of unified security strategies. Hospital IT environments consist of a patchwork of different software and hardware solutions, from Electronic Health Record (EHR) systems to specialized scheduling and billing platforms. Integrating these systems to work seamlessly is difficult, and the complexity creates opportunities for security vulnerabilities. Many hospital IT systems are built on outdated technologies that are difficult to upgrade or replace due to their critical role in maintaining patient care. As patients need treatment around the clock, IT systems must be available 24/7 to support critical functions, from patient care to administrative operations. As an example, HIS is deeply integrated with multiple other hospital systems, making upgrades and security patches extremely challenging. Medical devices rely on this system for essential functions, such as transmitting or sharing health information with other diagnostic equipment. This level of interconnectedness complicates implementing preventive measures, such as network isolation, which otherwise would be an effective strategy for threat containment.

In contrast to IT, MT and VT systems have traditionally been considered to be low-risk being operated in isolation without network connections. The devices in these domains have typically been managed by clinical engineers or employees of the technical facility management, who specialize in their operation and maintenance. The historical division persists today and operational expertise that is mostly beyond the scope of IT staff is still required. As an increasing number of devices becomes network-capable two disadvantages emerge from a security perspective. Firstly, the security teams, who are primarily integrated into the IT department, are separated from the engineers. Security frameworks and best practices are typically implemented in IT systems first, with other teams being informed later or overlooked entirely. Secondly, the separation comes along with limited observability. The increasingly interconnected MT and VT devices create blind spots concerning hospital's security infrastructure. Most often, security teams lack insight into what technology is running, making it impossible to detect suspicious activities. Many of the devices in MT and VT were not designed with modern security requirements in mind, relying on vendor-specific or outdated communication protocols and architectures. As a result, they frequently lack essential security features, such as encryption or effective access controls. The use of default credentials, which are rarely updated, is not uncommon and provides an easy entry point for malicious actors. The vendor-specific architectures impede the integration of standardized security measures. Unlike general IT systems, many MT and VT devices cannot easily adopt uniform security tools (such as a log forwarder to integrate these systems into a logging infrastructure) because the communication runs on proprietary systems or via proprietary protocols. Security patches and updates are entirely controlled by the manufacturers, who may be slow to convey updates. The VPN connections associated with remote vendor maintenance further restrict observability, as the

encrypted connections only terminate at or near the medical device. Therefore, these connections are barely monitorable for hospital SOC analysts. Even if security teams accomplish to monitor the connection, each vendor uses its own VPN solution, adding to the complexity. This diversity in implementations hampers the establishment of a standardized approach, demanding additional effort to monitor multiple systems effectively.

Beyond the requirements of both domains (MT and VT), the MT domain, with its variety of MCPS, introduces further specific characteristics that must be considered for the integration of SzA. Devices must comply with healthcare regulations, which frequently limit the types of modifications or updates that can be applied without undergoing a lengthy re-certification process. While these regulations are designed to ensure patient safety, they delay the implementation of security patches, leaving vulnerabilities unaddressed for extended periods. Moreover, the aforementioned vendor restrictions prevent the installation of third-party security software on devices, such as EDR agents, further restricting their ability to secure them effectively. As MCPS are often mission-critical, they cannot be easily taken offline for maintenance or security updates. Any security measure that introduces latency or requires downtime can disrupt patient care. The extended lifecycles of MCPS can result in the underlying software turning technologically obsolete, complicating efforts to apply necessary updates. Such unsupported systems represent a vulnerability, as any newly discovered weaknesses may never be patched, leaving them continuously exposed to threats. Additionally, security measures for MT must be carefully weighed with usability. If emergency access for authorized personnel is hindered or delayed due to security features, it can create significant challenges. For example, if a frequently used medical device always locks after a brief period of inactivity, this can lead to considerable disruption. Especially, if the unlocking process is dependent on a network connection, which might be unstable in certain areas of a hospital. In other domains and industries, it is common to keep access hurdles as high as possible. Contrary to this, in the MT domain, so-called "break-the-glass" access options exist. These options are intended to ensure that medical staff who would normally not have access to certain functions or data can still obtain them in an emergency. However, the balance between security and accessibility must be carefully controlled to avoid potential misuse.

All the discussed peculiarities make hospitals particularly susceptible to attacks and difficult to monitor for SzA. Existing vulnerabilities in one component within one domain can potentially compromise the entire network, leading to cascading effects. Conclusively, hospitals must prioritize and adjust existing frameworks for securing their digital infrastructures, ensuring comprehensive protection and timely response to incidents.

3.1.3 Scientific Advances in MCPS Attack Detection

Building on the assessment of the present state and the main challenges identified, this section outlines the *state of research and technology* in MCPS attack detection, aiming to refine the specification of *state of the art* for hospitals. For a more detailed review please refer to Chapter 4.

Overall, research efforts can be divided into two categories: applied and future-oriented research. Applied research addresses the industry's current challenges. The vast majority of researchers focus on improving anomaly-based approaches, e.g., by reducing false-positive rates. Research still pursues host-based and network-based approaches while focusing on insider threat detection. As existing challenges for MT at the host level also affect and restrict researchers, most rely on the network layer for attack detection. This layer provides the practical means to monitor multiple device types and communication protocols without directly facing all device-specific challenges. Researchers deciding on host-based detection approaches specifically address the current limitations of MCPS. For example, current work concentrates on using lightweight detection methods that account for resource constraints of medical devices (Thamilarasu et al. (2020), Meng et al. (2020)), or on developing privacy-preserving detection approaches (Almaiah et al., 2022). A promising research direction in this regard is the development of federated or transfer learning approaches, which enable attack detection systems to train models collaboratively across multiple distributed devices or data sources without the need to share raw data. Astillo et al. (2022), for example, present a novel approach where a submodel is generated on each host and subsequently fused to a central model on a cloud server. The minimized data exposure and reduced risks are a major step forward from traditional machine learning methods, often associated with considerable privacy concerns. Addressed by few is the topic of response automation. Here, attention is given to the challenge that an automated response could potentially endanger patients. For example, if an attack originates from a device that is still used for legitimate health-related operations, isolating it could impact the ability of a physician to control the device.

Future-oriented research comprises approaches that could enhance detection, e.g., by leveraging domain-specific properties of MCPS. Such domain-specific properties include incorporating health data into attack detection, e.g., correlating vital signs collected by different medical devices. If there are anomalies, an attack can be inferred (Gupta et al., 2021; Newaz et al., 2019). These research efforts show the industry-specific potential and offer a visionary glimpse into the future. While promising, those approaches are too recent and will not have a direct application in today's hospitals.

Independent of the difficulties encountered by hospital operators, it can be observed that research is confronted with its own distinct set of challenges. These challenges inherent to research may be a reason why the industry-specific problems have not yet been solved. The greatest challenge researchers currently face is the lack of standardized and comprehensive datasets specifically tailored to the medical field. To overcome this, some researchers have chosen to use datasets from other domains to validate their detection approaches, primarily analyzing non-domain specific network traffic data rather than patient information or healthcare-specific protocols (Hameed et al., 2021). Others generate synthetic data to simulate real-world medical environments (Mitchell and Chen, 2015). While these substitutions work for single approaches, they complicate cross-comparisons and hence impair further advancements in the field. The scarcity of datasets is hindering the development of effective detection methods for the variety of network protocols utilized exclusively in hospitals. Researchers need standardized environments to consistently test and benchmark their attack detection approaches. Developing

domain-specific, open datasets and collaborative testing environments is critical to advancing anomaly detection and machine learning models tailored to hospital networks.

Furthermore, certain issues challenge operators and researchers, leading them to rely on workarounds or assumptions, while the underlying questions remain unanswered. For example, there appears to be no consensus whether hospital network traffic is consistently encrypted. This uncertainty has resulted in varying conclusions and detection strategies among researchers. Some assume that a large portion of hospital network traffic is encrypted, directing them to propose flow-based detection approaches rather than payload inspection (Fernández Maimó et al., 2019). Conversely, others extract patient data from captured network traffic, assuming that such data is transmitted unencrypted (Hady et al., 2020).

Although intensive research is being carried out in the field of MCPS attack detection, basic challenges are still unsolved. Already Clark and Fu (2012) identified two challenges related to medical device security: "(1) computer security researchers seldom have access to real medical devices for experimentation, and (2) the computer security community is largely disjoint from the biomedical engineering community." These challenges persist to this day and, therefore, may explain why comprehensive attack detection for MCPS is not yet *state of the art*.

3.1.4 Determining the Gap Between Research, Compliance, and Practice

This section brings together the present state challenges and research findings to determine the gap between the current and foreseeable capabilities of hospital operators and the regulatory requirements of the IT Security Act 2.0. The challenges identified highlight that certain areas are unable to effectively implement the measures outlined by the OH SzA. The following discussion examines these discrepancies to investigate how the many fallback options offered by the OH SzA represent solutions for the IT, MT, and VT domains.

Components of the digital infrastructures of hospitals that are also used in other sectors and industries can be effectively monitored with the help of prevalent attack detection solutions. Here, integration blueprints into SIEM systems exist, so that general recommendations given by the OH SzA are applicable to domains such as the general administrative IT. Specialized hospital IT systems, like HIS and PACS, are typically operated on common IT servers as well. Consequently, existing blueprints can generally be applied at least at the OS level. However, monitoring the industry-specific applications running on these systems may require certain modifications to account for their unique operational and security characteristics.

In contrast to the IT domain, traditional detection systems such as EDR are often entirely unsuitable for MT and VT, as those devices may lack the necessary computing power, or have strict certification requirements so that no agents may be installed. As first fallback provided by the OH SzA, these systems could use existing log forwarding capabilities to be connected separately to a SIEM. Nevertheless, this option may fail to be feasible for all MT and VT, as they are managed by manufacturers frequently so that hospital operators have no or only restricted access to the systems. The long lifetime of these devices entails that this will not change in the

foreseeable future. The second fallback option provided by the OH SzA focuses on monitoring at the network level. The positive trend towards a higher percentage of encrypted network traffic may cause challenges for logging approaches. Limited or no access to MT results in unretrievable certificates and traffic can therefore not be decrypted for in-depth analysis. This is compounded by the various VPN solutions used by manufacturers for remote maintenance. Additionally, while network-level monitoring offers the benefit of covering multiple devices, it also comes with the disadvantage of an increased false-positive rate for incidents (Knerler et al., 2023). The assessment of the *state of research and technology* in particular showed that although the network level is intensively studied by scientists, it is currently only moderately promising. This is mainly due to the fact that cooperative security research in the field of MT has been neglected for decades. A large number of protocols prevalent in hospitals that need to be integrated into detection approaches face a lack of training datasets. This results in a limited protocol understanding so that an evaluation of network traffic will remain restricted to non-industry-specific protocols in the near future. Therefore, solely metadata or network flow data can be evaluated for industry-specific protocols. Here, many detection mechanisms are constrained to identifying known threats (e.g., by using signatures or basic parameters like IP lists) or simple anomalies. As discussed in Chapter 1, more and more attackers are directly targeting hospitals and might use proprietary communication protocols, which means these basic metrics are insufficient. The final fallback offered by the OH SzA is to take unspecified alternatives to protect devices for which all other measures are not sufficient. Since there are no blueprints for such alternative measures, it is up to the B3S to specify them.

The OH SzA also provides fallback options in the area of automated response. Here, the present-state assessment uncovers that hospital operators are very cautious. Research highlights valid reasons for this, as there are no mitigations for the risks posed by automated intervention in the area of the kDL. The risks outweigh the potential. In general, the OH SzA demands automated intervention. However, the first fallback already takes the aforementioned risks into account and hence limits the need for automation in this respect. For the specification of the *state of the art*, automated interventions should therefore only be recommended outside the kDL. Nevertheless, automated response mechanisms are not limited to an all-or-nothing approach. They offer a spectrum of options that can be tailored to specific scenarios. Automation should be used to support security analysts during an incident, for example, in the form of alert enrichment. This provides analysts with faster and more comprehensive information and therefore enables them to act more quickly.

In conclusion, the requirements and fallback options outlined in the OH SzA can be applied to differing extents across the domains of IT, MT, and VT. Their applicability and effectiveness vary depending on the types of devices. Furthermore, some fallback options are only contoured but not described in detail. Therefore, in the following sections, suggestions are discussed for alternative measures and, thereby, for a specification of the industry-specific *state of the art*.

3.2 The Potential of Honeypots in the Medical Domain

Honeypots function as decoy systems designed to lure attackers into interacting with them, thereby capturing data on malicious behavior without exposing real systems. In IT environments, honeypots are deployed to mimic legitimate systems and collect information on attackers' techniques and tools. This section introduces honeypots and explores how they might offer solutions to some of the limitations identified in current attack detection in hospital environments.

3.2.1 Addressing Detection Challenges with Honeypots

The concept of honeypots as a tool for cyber security was first popularized in Stoll (1989). Clifford Stoll described how he used a rudimentary honeypot to track an intruder who had breached his laboratory's computer systems. By creating a controlled environment that appeared to be a valuable target, Stoll was able to deceive the attacker and observe the activities, gaining insight into tactics and motivations. His efforts highlight the effectiveness of deception as a strategy for detecting and studying unauthorized access, marking an early practical application of honeypot technology.

The Deception Toolkit (DTK), introduced by Cohen (1999), is a freely available tool designed to emulate known vulnerabilities to lure attackers. The DTK represented an evolution in honeypot strategy, transitioning from merely observing attackers to actively misleading and deceiving them, thereby integrating the concept of proactive defense. This tool was specifically designed to mimic common services and servers, enabling companies to detect and log unauthorized access attempts. This illustrates the utility of honeypots for commercial use cases beyond research or experimentation. The goal was not necessarily to gather intelligence on attackers' behaviors but to create an illusion of vulnerabilities that would deter attackers by making their presence known. Since these early implementations, honeypots have continued to evolve and are now extensively used in industry and research to gain insights into emerging attack techniques (Franco et al., 2021). By deploying honeypots, researchers can gather real-time data on the latest attacker methodologies, tools, and motives, enabling a deeper understanding of threat landscapes and improving the development of more robust security measures.

Honeypots can broadly be categorized into three types: low-interaction, medium-interaction, and high-interaction honeypots. Low-interaction honeypots simulate a limited set of services and provide minimal interaction with attackers. They are relatively easy to deploy and maintain. Their primary purpose is to detect unauthorized access attempts and log basic attack patterns. Medium-interaction honeypots provide a greater degree of interaction compared to low-interaction types, often emulating specific services more convincingly without offering full system capabilities. These honeypots allow attackers to perform limited actions, giving defenders more insight into the methods being used while still maintaining a controlled environment. High-interaction honeypots, on the other hand, are the most sophisticated type and involve real operating systems and services. They enable attackers to engage fully with what appears

to be a legitimate system. While high-interaction honeypots provide the most detailed data on attacker behavior, they are also more resource-intensive to maintain and can pose significant risks if an attacker successfully compromises a honeypot and uses it as a staging point for further attacks (Franco et al., 2021).

Attackers have kept pace with these developments and are aware of the potential presence of honeypots in target environments. Many attackers now use techniques to identify whether a system is a honeypot before launching a full attack, such as probing for inconsistencies that might reveal a deceptive system (Abbas-Escribano and Debar, 2023). This has led to an ongoing competition in which honeypot designs must continually evolve to become as realistic as possible. To remain effective, honeypots need to mimic real systems with high fidelity, from network responses to system artifacts, ensuring that they do not give away any clues that could reveal their true nature. The challenge lies in making honeypots indistinguishable from legitimate systems, thereby increasing the likelihood that attackers will engage with them and allowing defenders to gather the intelligence needed to improve their security posture.

Honeypots hold great potential for addressing the current challenges in attack detection for hospital networks and offer several distinct benefits. Since honeypots operate as isolated systems, they do not interfere with the actual functioning of critical systems, making them an attractive option for hospital environments. One of their main advantages is that they are not bound by regulatory or certification requirements that would hinder their integration into existing detection infrastructures. Honeypots can be deployed freely within meaningful network segments, enabling targeted monitoring of potential attack vectors without the constraints of compliance. As a result, alerts triggered by honeypots, such as attacks on medical device emulations, can be sent directly and reliably to analysts for investigation. This improves observability and detection efficiency for security operations. From a scientific perspective, high-interaction medical honeypots could also provide valuable insights. Deploying these advanced honeypots would allow researchers to assess the degree to which medical technologies are targeted by attackers and analyze the specific attack methods used. This type of intelligence could be used to develop better defenses and enhance the overall resilience of healthcare infrastructures.

Currently, there are only a few publicly available, domain-specific honeypots for the hospital industry. Two well-known examples are DICOM-Pot (Keri, 2023), which focuses on the Digital Imaging and Communications in Medicine (DICOM) protocol and MedPot (Schmall, 2022), which simulates communication based on the Health Level 7 (HL7) standard. HL7 and DICOM are fundamental protocols to hospital networks. HL7 facilitates the exchange of clinical and administrative data, whereas DICOM is responsible for managing and transmitting medical imaging data. Emulating these protocols in honeypots provides valuable insights into how attackers may exploit healthcare-specific communication patterns. However, DICOM-Pot and MedPot are low to medium-interaction honeypots. They do not replicate specific MT with a high degree of realism, nor do they offer extensive configurability. To understand the limitations, it is worth considering the variety and exposure of the protocols. Forescout (2024b) reports a growing number of DICOM ports exposed to the public Internet, with Germany among the top three countries with the highest prevalence of such Internet-facing devices. In

the heterogeneous MT ecosystem the DICOM protocol is associated with one of the top ten open TCP ports. According to Forescout (2024a), more than 40 different categories of medical devices are observed to expose DICOM ports. Yet, because DICOM-Pot is designed to mimic only a generic PACS, it cannot be configured or adjusted to realistically represent the full range of imaging-related devices and their associated security risks. A similar challenge arises with MedPot. While it employs the latest iteration of the HL7 standard – Fast Healthcare Interoperability Resources (HL7 FHIR) – the version has not yet been widely adopted in everyday hospital operations. Although HL7 FHIR was introduced in 2014, its predecessors, particularly HL7v2, remained dominant well beyond 2018 (Joyia et al., 2018). Even in 2024, a recent survey by Health Level Seven International (2024) found that most respondents use HL7 FHIR for isolated use cases only. Although 84% anticipate broader adoption in the coming years, the slow uptake suggests that a honeypot relying on HL7 FHIR may not accurately reflect typical hospital environments or represent an attractive target for adversaries. Instead, devices employing older versions like HL7v2 and HL7v3 – which are still widely in use – would likely offer more tempting targets for attackers.

In summary, the heterogeneity of medical devices in hospital environments cannot be adequately represented by the currently available, domain-specific honeypots. Hospitals utilize a wide array of medical equipment, each with different communication protocols and usage patterns, which impedes a single, low- to medium-interaction honeypot to comprehensively represent these devices. The presented honeypots are not easily adaptable, which limits their ability to realistically emulate specific types of medical technology. This prevents them from being tailored to the nuances of different devices, which is essential for effective deception. In addition, these honeypots are easily recognizable. Attackers can identify the discrepancies between a honeypot and a legitimate medical device, especially if the honeypots provide limited functionality and restricted access to protect against potential compromises. This reduces the value of the honeypot as a tool for understanding and mitigating real-world threats. Addressing these challenges requires researchers to develop more flexible and high-fidelity honeypots capable of mimicking the diversity of medical devices and networks found in healthcare environments.

3.2.2 Advancing Honeypots with Large Language Models

While the current challenges of honeypots for hospital systems can't be addressed without substantial research and development effort, a new research direction has the potential to streamline the process: LLM-based honeypots.

LLMs represent an emerging area of artificial intelligence and are capable of contextualizing conversations, producing source code, and generating other forms of machine output (Biswas, 2023; Ray, 2023; Zhang and Li, 2021). These models have the ability to understand context and generate a wide variety of responses, allowing them to adapt effectively to various scenarios. As server attacks can resemble technical conversations where attackers send commands and expect specific responses, the capabilities of LLMs could be particularly well-suited to

handle such interactions. By interpreting incoming commands and generating contextually appropriate responses, LLMs could effectively replicate interactions taking place between attackers and actual medical devices or other hospital systems. This flexibility positions LLMs as a promising candidate to address key limitations of current honeypots – in particular, the high development costs and the tendency for attackers to easily identify traditional honeypots. Furthermore, using LLM-based honeypots would reduce the risk of unintentional breakouts. Traditional honeypots carry the risk of being exploited by attackers as a staging ground to infiltrate further systems, which poses challenges for their deployment in production environments. LLMs, by generating responses without executing real system commands, inherently reduce the likelihood of such an occurrence, enhancing their suitability for secure environments. Compared to other domains, the problem of LLMs "hallucinating" and generating convincing but factually incorrect responses might be less critical in the context of honeypots. For honeypots, the primary goal is to convince the attackers that they are interacting with a genuine system, rather than focusing on the technical accuracy of every response. The ability of LLMs to generate persuasive, albeit not entirely correct, responses can effectively deceive attackers and fulfill the intended defensive purpose.

Several research efforts have already highlighted the potential of LLMs across a wide range of protocols and environments, including various operating systems (Windows, Linux, macOS (McKee and Noever, 2023)), MySQL servers (Hu et al., 2024), IoT devices (Mfogo et al., 2023), Modbus, and S7comm (Vasilatos et al., 2024). These studies illustrate the versatility of LLMs in emulating numerous types of systems and protocols, thereby suggesting their potential effectiveness in designing more sophisticated honeypots for medical environments. Nevertheless, existing research in this field is not easily comparable, as each study evaluates performance differently (Weber et al., 2024b). While the diverse methods are useful for demonstrating the broad potential of LLM-based honeypots, they are insufficient for advancing the research field in a systematic way, as the results cannot be compared to one another. To address this, Chapter 5 proposes a novel, unified evaluation approach based on paraphrase-mining. This approach seeks to provide a consistent framework for assessing the effectiveness of LLM-based honeypots, fostering advance in the field towards practical application.

3.3 Leveraging MDS2 for Security Evaluations

As explained in Section 3.1, significant uncertainties among hospital operators and researchers regarding the security features of medical devices exist, particularly in the context of MCPS. From the operators' perspective, these uncertainties often revolve around the visibility of risks associated with MCPS and how to ensure the safe integration and operation of those devices within healthcare environments. Researchers, on the other hand, are focused on understanding the inherent security properties of the devices to develop and validate security models. This thesis shows that MDS2 documents offer valuable insights into the security capabilities of medical devices. Furthermore, it investigates their potential to systematically assess the overarching security properties of MCPS, to answer the open questions of operators and researchers.

3.3.1 The Current Role of MDS2 in Hospital Security

The Manufacturer Disclosure Statement for Medical Device Security was developed to provide a standardized and structured method for manufacturers to disclose security features of their products. Initially conceived as a questionnaire comprising 43 questions, MDS2 aims to address fundamental concerns regarding the security properties of medical devices, such as their ability to protect patient data and ensure device integrity. Over time, MDS2 evolved into a more comprehensive framework, and the latest version, released in 2019, includes 216 questions organized across 23 security-related categories. This evolution reflects the growing complexity of medical devices and the need for more detailed security transparency in healthcare environments. Now, MDS2 documents hold essential information tailored specifically to MCPS security. They encompass details about authentication methods, data encryption capabilities, logging mechanisms, and other key security features implemented within a device. The increasing international adoption of MDS2 by manufacturers and healthcare institutions has further enhanced its relevance, establishing it as a tool to improve the security posture of medical technology on a global scale. Today, MDS2 documents are used by healthcare organizations to assess the IT security of single medical devices they have in use, without the need of direct interaction with those devices. With these documents, healthcare professionals can quickly and effectively evaluate the security capabilities of a medical device, helping them to make informed decisions regarding device deployment and ensuring that security risks are adequately mitigated. Thereby, the role of MDS2 in enhancing transparency between manufacturers and device operators strengthens the overall security within hospital environments.

3.3.2 Expanding the Use of MDS2 for Security Landscape Analyses

Chapter 6 proposes the first systematic assessment of MDS2 documents, highlighting a valuable opportunity for these documents to be used beyond their original purpose. They can serve as a powerful tool for deriving a comprehensive security landscape of medical devices, benefiting researchers and operators. Such systematic usage of MDS2 data not only fosters a deeper understanding of the current security landscape but also supports proactive security management in an increasingly connected healthcare environment.

From a research perspective, MDS2 documents offer a wealth of information that, when aggregated and analyzed, can yield insights into the broader security posture of medical devices. This thesis is the first to determine how many medical devices are capable of encrypting network traffic, providing a quantitative perspective on the prevalence of encryption capabilities in medical technology. Additionally, it identifies how many devices log security-relevant events and, more importantly, how many of those devices are capable of forwarding such logs to a SIEM system. This is critical, as on-device logging capabilities alone are not sufficient for proactive defense. Researchers can use this information to further develop and refine detection approaches for MCPS, certain of whether they can expect and use log data or decrypted network traffic. Furthermore, an explanation for earlier findings that medical devices are particularly susceptible to attacks involving standard passwords is discovered. The analysis of

MDS2 documents revealed that a substantial number of medical devices lack the capability to modify default passwords, explaining why these vulnerabilities persist and emphasizing the need for manufacturers to address this issue.

The systematic assessment of MDS2 documents also holds potential value for healthcare operators by enabling them to generate IT security status reports specific to their organizations. Selecting only the information pertaining to devices in use within their environment, operators could quickly compile an overview of the current security posture of their medical device inventory. Such reports can help healthcare organizations to prioritize risk mitigation efforts, allocate resources more efficiently, and fulfill regulatory compliance requirements. The information about logging capabilities allows operators to determine which medical devices can be directly integrated into a detection infrastructure and the information about encrypted network traffic helps to decide if an NIDS might be beneficial in a network segment containing MCPS. Additionally, MDS2 documents can be leveraged to identify which devices are particularly vulnerable due to the lack of certain security features, such as unnecessary open ports and services not being disabled. This knowledge allows healthcare organizations to target specific devices that require additional security measures, reducing their overall risk exposure. Many MDS2 documents also include a list of third-party libraries and software components in the form of Software Bill of Materials (SBoM). Aggregating these SBoMs and analyzing them helps hospital security teams rapidly assess whether a discovered vulnerability in a specific software component impacts their organization. By identifying which devices are affected, security teams can quickly determine which medical devices require patches or other mitigating actions, streamlining the vulnerability management process.

3.4 Synthesis: A Way Forward for Hospital Security

The preceding sections highlighted the gaps between research, compliance, and practice, revealing substantial challenges in securing specialized IT, MT, and VT in hospitals. While regulatory requirements such as the OH SzA establish a baseline for attack detection, hospitals face unique constraints that limit the adoption of such general good practices. Synthesizing these insights proposes a forward-looking, multi-faceted approach to bridge existing gaps and advance hospital security.

In accordance with the Handbook of Legal Formality the assessment of the present state of SzA in hospitals and the advancements in research allow that conclusions can be drawn to specify the *state of the art*. This aids in identifying alternative and more feasible methods and provides a more precise specification. By reviewing and orienting towards general standards and good practices for attack detection, the applicability of recommended measures can be estimated. Technologies such as SIEM, EDR, and XDR can be seen as legal obligations under the IT Security Act 2.0 in combination with the OH SzA and need to be employed where sensible. While these tools are integral to achieving compliance, as previously discussed, compliance alone is not sufficient to guarantee security. Digital hospital infrastructures are marked by complexity, particularly regarding MT and VT components, which are characterized by a

diversity of devices, certification-related requirements, and resource limitations. Certifications not only hamper the installation of EDR/XDR agents in the area of MCPS, but also the decryption and inspection of network traffic, as corresponding certificates may not be adapted, modified, or retrieved. Additionally, industry-specific protocols are rarely supported by available NIDS solutions, even if they are not encrypted. Due to the long lifespans, proprietary systems, and inherent limitations, all these challenges will persist for the foreseeable future. Therefore, all but the last fallback option implicitly provided by the OH SzA are often infeasible. A gap remains between what hospitals are currently able to implement and comprehensive attack detection. This gap arises primarily from fundamental, industry-specific challenges. A more nuanced approach to security is necessary, exceeding basic regulatory adherence and incorporating solutions tailored to the specific needs of healthcare systems.

Honeypots, potentially facilitated by LLMs, present promising opportunities for tackling the unique challenges of detection in MT and VT environments. They offer a solution to the challenges mentioned above by emulating these devices, attracting attackers, and allowing hospitals to monitor for threats without requiring direct interventions on the devices themselves. Consequently, they might be a suitable alternative in reference to the last fallback option of the OH SzA. This dissertation marks a critical step forward in adapting detection capabilities to meet the specific requirements of MT and VT. The use of LLMs to enhance the scalability and adaptability of honeypots can enable hospitals to deploy dynamic detection systems that are more responsive to the evolving threat landscape. However, at the time of writing, honeypots are not ready for immediate deployment and are more realistically expected to become viable in the mid-term.

Until then, hospitals can leverage MDS2 as an actionable tool. By systematically evaluating the security readiness of their medical devices using MDS2, hospitals can identify those medical devices that can be integrated into a traditional detection infrastructure. MDS2 can also help to prioritize which devices receive the security analysts' attention next in order to configure them more securely or remediate weaknesses. Furthermore, they can support undertaking informed decisions on topics such as network segmentation and monitoring. MDS2 documents offer a clear and structured path for incremental improvements in device security, providing hospitals with the information needed to enhance their overall IT security framework step by step.

Ultimately, the future of secure digital hospital infrastructures lies in the ability to merge practical, short-term measures such as leveraging MDS2, with forward-looking research into advanced detection methods. By closing the gap between current capabilities and scientific innovation, hospitals can develop a comprehensive security strategy that meets today's regulatory requirements and tomorrow's security challenges. A cohesive approach will ensure that hospitals are better equipped to safeguard patient data and maintain operational integrity in an increasingly connected and vulnerable environment so that patients' lives can be protected in the long term.

Chapter 4

Attack Detection for Medical Cyber-Physical Systems – A Systematic Literature Review

This chapter gives an overview of the contributions and the impact of **Weber et al. (2023)**[1]:

Simon B. Weber, Stefan Stein, Michael Pilgermann, Thomas Schrader

"Attack Detection for Medical Cyber-Physical Systems — A Systematic Literature Review"

In: *IEEE Access*, Volume 11, pages 41796-41815, April 2023

Acceptance Rate: ~27%

4.1 Summary

This paper explores the landscape of research on detecting attacks in MCPS. Given the critical vulnerabilities within hospitals and the interconnected nature of medical devices, this review emphasizes the unique challenges in securing these systems, including the diversity of device types, connectivity methods, and specificities of medical terminology. The authors conducted a comprehensive review following the guidelines of Kitchenham, Charters, et al., 2007 by formulating and answering six research questions focused on the methods, data sources, and types of threats addressed in the current literature. Thereupon, five future research topics were derived.

One key observation of the study is the prevalent use of anomaly-based detection approaches. These methods are favored for their capacity to identify deviations from typical network behavior, though they frequently struggle with high rates of false positives. The reviewed works tend to prioritize the detection of insider threats over external attacks. Detection on the network level is preferred due to the absence of standardized, accessible data from individual medical devices. Device-level data collection is fragmented, often tied to specific devices, and lacks uniformity, complicating the application of solutions on a broader scale. Publicly available data

[1] ©2023 IEEE. Reprinted, with permission, from Weber, S. B., Stein, S., Pilgermann, M., & Schrader, T. "Attack detection for medical cyber-physical systems–a systematic literature review". IEEE Access. (2023).

for MCPS-specific traffic and attacks is scarce, prompting researchers to generate their own datasets or use custom testbeds to mimic real-world conditions. This scarcity hinders comparability between studies and limits the applicability of many solutions to actual healthcare environments. Moreover, the study calls attention to the need for approaches that incorporate medical context with technical indicators to improve the accuracy and relevance of intrusion detection in hospitals.

In response to the challenges identified, five future research directions were proposed. These suggest capturing the distinct technical requirements of healthcare networks, developing MCPS-specific datasets, advancing standardization of device communication protocols, integrating medical data with traditional technical indicators, and adopting risk-aware prevention methods. The latter must consider patient safety when implementing automated preventive measures, balancing the need for security with the critical importance of uninterrupted patient care. This review underscores the urgency to develop robust and adaptable attack detection systems tailored to the healthcare sector to ensure patient safety and protect sensitive data in an increasingly complex threat landscape.

4.2 Personal Contribution

Simon Weber designed the concept of the study, conducted the database search, and retrieved, organized, and prepared the documents. Following the guidelines for structured literature reviews in software engineering by Kitchenham, Charters, et al., 2007, the screening and selection of the papers was performed independently by Simon Weber and Stefan Stein. For the subsequent data extraction, Simon Weber served as the data extractor and Stefan Stein as the data checker. Any discrepancies in final classifications were discussed collaboratively among all authors until consensus was achieved. Simon Weber authored the manuscript. Stefan Stein, Michael Pilgermann, and Thomas Schrader reviewed the drafts and provided feedback.

4.3 Importance and Impact on this Thesis

As discussed in section 3.1 and aligned with the Handbook of Legal Formality (Bundesministerium der Justiz, 2024), the *state of the art* ranges between the *generally accepted rules of technology* and the cutting edge of scientific advancement. Understanding the latest research developments is, therefore, crucial for formulating industry-specific *state of the art* specifications. By identifying the current challenges faced by researchers, this review helps to anticipate which technologies might become available in the mid-term. Given the limitations in recent attack detection technologies as well as in current research endeavors in the context of addressing specialized devices in the healthcare industry, it is clear from this review that supplementary measures are required. Furthermore, the insights demonstrate the importance of addressing researchers' unresolved questions, enabling future studies to establish accurate assumptions that can guide advancements in MCPS attack detection.

Attack Detection for Medical Cyber-Physical Systems - A Systematic Literature Review

SIMON B. WEBER
Heinrich-Heine-University
Düsseldorf, Germany
Simon.Weber@hhu.de

STEFAN STEIN
University of Applied Sciences
Brandenburg a. d. Havel, Germany
Stefan.Stein@th-brandenburg.de

MICHAEL PILGERMANN
University of Applied Sciences
Brandenburg a. d. Havel, Germany
Michael.Pilgermann@th-brandenburg.de

THOMAS SCHRADER
University of Applied Sciences
Brandenburg a. d. Havel, Germany
Thomas.Schrader@th-brandenburg.de

Abstract

The threat situation due to cyber attacks in hospitals is emerging and patient life is at risk. One significant source of potential vulnerabilities is medical cyber-physical systems (MCPS). Detecting intrusions in this environment faces challenges different from other domains, mainly due to the heterogeneity of devices, the diversity of connectivity types, and the variety of terminology. To summarize existing results, we conducted a structured literature review (SLR) following the guidelines of Kitchenham et al. for SLRs in software engineering. We developed six research questions regarding detection approach, detection location, included features, adversarial focus, utilized datasets, and intrusion prevention. We identified that most researchers focused on an anomaly-based detection approach at the network layer. The primary focus was on the detection of malicious insiders. While several researchers used publicly available datasets for training and testing their algorithms, the lack of suitable datasets resulted in the development of testbeds consisting of various medical devices. Based on the results, we formulated five future research topics. First, the special conditions of hospital networks, the MCPS deployed within them, and the contrasts to other IT and OT environments should be examined. Thereupon, MCPS-specific datasets should be created that allow researchers to address the health domain's unique requirements and possibilities. At the same time, endeavors aimed at standardization in this area should be supported and expanded. Moreover, the use of medical context for attack detection should be further explored. Last but not least, efforts for MCPS-tailored intrusion prevention should be intensified. This way, the emerging threat landscape can be addressed, IT security in hospitals can be improved, and patient health can be protected.

Index Terms

Detection, IDS, Intrusion Prevention, Medical Cyber-Physical Systems, Medical CPS, Internet of Health Things, IoMT, Medical IoT, Connected Health, Healthcare 4.0

1 Introduction

The healthcare sector faces an increasing threat of cyber attacks. A Comparitech study explored the threat landscape of the US sector and found that in the time 2016 to 2022, 6,835 healthcare companies were hit by ransomware [1]. Already in 2014, a SANS report admonished the risks of MCPS and identified "Nontraditional medical endpoints" as one of the main malicious traffic sources [2]. Gartner predicts for the year 2025 that operational technology (OT) environments will be weaponized to harm or kill people and that the resulting financial impact from such attacks will amount to $50 billion per year [3].

Coventry et al. surveyed hospital staff to determine the reasons for clinics' high IT security risks. One key finding is that medical device software is often outdated and unsupported [4]. This corresponds to the report of the European Union Agency for Network and Information Security (ENISA). They stress that legacy software and unpatched vulnerabilities are particularly critical in the healthcare sector. Accordingly, imaging systems, patient monitoring, and medical device gateways root for 86% of hospital security issues [5]. As Coventry et al. emphasized, securing legacy medical devices seems more crucial than ever. Despite these findings, current security scanners often fail to detect vulnerabilities in the health-

care environment because they do not have modules for medical devices and systems [6].

Intrusion detection is a technology that has been around for more than three decades [7]. Many organizations rely on it even more in a time of increasing cyber threats. Its most widespread application is in the field of (office) information technology (IT). In contrast, the used OT is not as covered in many sectors. Only in recent years, there have been efforts to transfer insights from IT to OT, primarily because of the emerging threat situation [8]. Many sectors can interoperate and share their sector-specific discoveries and perceptions. The health sector differs in this regard. There are three reasons: The heterogeneity of OT devices in healthcare, the diversity of connectivity types, and the variety of terminology. The heterogeneity of devices and missing regulations lead to a situation where no central management of devices from assorted manufacturers is possible. Furthermore, the different needs of different devices lead to different requirements for connectivity. E.g., while computed tomography scanners are regularly connected via a wired connection, wearable medical devices require a wireless connection for obvious reasons. Researchers cover this domain as wireless sensor networks (WSN), of which subgroups are medical smartphone networks (MSN) and wireless body area networks (WBAN). These networks consist of devices known as wearables, which are worn by a person and can also be connected to each other. Since devices of those groups are carried around, they not only have special requirements for connectivity but also for an intrusion detection system (IDS). In addition, the need for real-time detection for all MCPS, regardless of the device type, is argued to be even more critical than in existing mechanisms because lives could depend on a timely detection [9]. While some researchers try to detect intrusions in a protocol-agnostic way, others are motivated by the particular conditions of a subset of medical devices. This leads to different categorizations and definitions of groups and, thereby, to various terms. We further discuss the diversity of terms in section 3.2.

This paper aims to outline the existing research on attack detection in the healthcare sector. We focus on the current state of research in detecting attacks on medical devices available for hospitals and clinics. The challenges of hospital networks, attached medical devices, and the plethora of protocols used by those devices are of particular interest. By discussing this environment's background and special requirements, we identify research gaps and give future endeavors direction.

The paper is organized as follows. In section 2, we present other secondary studies and point out how this paper complements the existing work. Thereafter, in section 3, we describe our methodology and present the research questions, according to which we have evaluated the studies. The results are presented in section 4. In section 5, we discuss the findings and work out the implications. Finally, we draw a conclusion in section 6.

2 Related Work

Existing reviews and survey papers on MCPS attack detection can be summarized into four groups. The papers of the first group discuss work about general IoT and merely touch the area of medical devices. They refer to MCPS either as motivation or to highlight them as a unique area with particular characteristics. One example is Banerjee et al., who discuss the security of several sectors in which IoT is used and how blockchain could improve it in the future. The healthcare domain is a characteristic example in which very sensitive data must be shared, and privacy is essential [10].

The second group concentrates on medical device security and attempts a comprehensive overview. Either they use broad definitions of security and include not only IT security but also privacy and patient safety, or the review outlines several IT security measures. Examples are Yaacoub et al., Tervoort et al., and Ferrag et al., who provide a detailed overview of relevant attack scenarios for medical devices and discuss which defensive measures can protect the devices from which attacks. These measures range from technical to non-technical aspects. Yaacoub et al. recommend a layered security architecture, ranging from raising awareness through employee training to sophisticated intrusion detection, mainly through a machine learning (ML)-based intrusion detection and prevention system cooperating with honeypots and security information and event management (SIEM) to gain the latest insights into attacks [11]. Tervoort et al. conduct a scoping review presenting an overview of security solutions for medical software vulnerabilities that do not require the software to be replaced. Besides intrusion detection, monitoring specific aspects of medical devices, such as software execution characteristics and tunneling legacy protocols, have been examined [12]. Ferrag et al. outline security solutions for the Internet of medical things (IoMT) of five categories: authentication and access control, key management and cryptography, intrusion detection systems, blockchain-based solutions, and privacy-preserving solutions [13].

The third group of papers focuses on a specific aspect or a specific type of approach. Thomasian and Adashi summarize the policy and regulatory measures (primarily concentrated on the US) to secure medical devices. Furthermore, they provide an overview of the emerging threats in this context on a high level [14]. Hameed et al. elucidate ML-based approaches in their structured review of security and privacy in the context of the IoMT. Besides insightful statistical data surrounding the publications, such as the geographical distribution of research groups and the development of publications per year, a focal point of the work of Hameed et al. is ML-based intrusion detection [15]. Rbah et al. concentrate their efforts on comparing deep learning methods utilized for IDSs in the IoMT. They observe that many researchers develop their approaches in an isolated environment for a limited number of attacks [16]. Pelekoudas-Oikonomou et al. review blockchain-based security mechanisms for IoMT edge networks. While they describe several ways in which attack detection in IoMT-edge networks could benefit from a blockchain extension, they state that, to their

knowledge, there are no blockchain-based IDSs specifically designed for IoMT-edge networks yet. Instead, they outline approaches from other IoT environments and show how they could be applied to IoMT-edge networks [17].

The fourth group of papers compares approaches tailored to small subsets of MCPS. Eliash et al. discuss the security of the subset of medical devices used in intensive care units (ICUMDs), introduce a taxonomy for these devices, and explain how these devices interact with each other. They develop scenarios for 16 attacks on medical devices and derive the main building blocks. Additionally, they analyze the applicability of existing security mechanisms, including detection mechanisms [18]. Similarly, Kintzlinger and Nissim establish a taxonomy for personal medical devices (PMDs) and collect attack scenarios and building blocks for attacks on this group of devices. Furthermore, they review the existing security solutions and identify the gaps between them and the identified attack vectors [19]. Ghosal et al. present a survey for ML approaches utilized for IT security in cloud-based IoT healthcare systems [20], Wa Umba et al. review security measurements exclusively for software-defined WSNs (SDWSNs) [21], and Wazid et al. compare detection approaches for malware in the IoMT environment [22].

This paper differs from those presented as it aims to provide an overview of all intrusion detection approaches for all kinds of medical devices available for hospitals and clinics. The main contributions can be summarized as follows:

- We present the current state of research on attack detection in medical cyber-physical system environments. In particular, we show the various challenges that are special or unique to the health sector and frame our research questions around these specifics.
- As a distinct difference from other secondary studies based on a single or small number of keywords (e.g., IoMT), we identified 22 synonyms for MCPS. We included them in an extensive database search as a basis. The high number of synonyms allows a comprehensive and profound analysis of the research state.
- By following the guidelines of Kitchenham et al. for structured literature reviews in software engineering, we minimized the risk of a biased consideration of the studies available. This includes:
  1. A structured two-step screening process
  2. Transparent inclusion and exclusion criteria for study selection
  3. The independent review of studies by at least two researchers in every selection and extraction step.
- By answering six research questions, we structure the confusing and convoluting state of literature and highlight commonalities and differences. For exceptional approaches, we present a detailed description.
- We critically engage with the selected aspects of the research and discuss the applicability of the proposed approaches.
- The resulting discrepancies will help researchers conduct more focused research through five derived future research topics.

3 Methodology

We adopted the guidelines for performing systematic literature reviews (SLR) in software engineering [23]. According to Kitchenham et al., the goal of such a review is threefold: Firstly, the review shall summarise the existing results in a field. Secondly, it should identify gaps in the current research, and thirdly, it should provide the background to position future research endeavors.

3.1 Research Questions

We developed six research questions to determine the state of research in the field of MCPS attack detection. These are outlined in the following.

Which detection approach is used?

First, we wanted to ascertain what detection approaches are utilized most to detect attacks in hospital environments. Research knows three types of IDSs:

- signature-based detection
- anomaly-based detection
- specification-based detection

The two best-known subcategories are signature-based and anomaly-based IDSs. Signature-based IDSs use predefined patterns of known attacks to detect intrusions in a pattern-matching approach. The major downside is that those systems can only detect known attacks. Even the smallest changes that modify the signature of the attack might evade detection. The upside is few false alarms.

The counterpart is anomaly-based intrusion detection which has drawn much interest in the research community. Those IDSs model the expected behavior of a system or network and warn in the case of deviation from baseline behavior. Advantages and disadvantages are contradictory: While this approach might detect even zero-day attacks, it is difficult to consider every borderline case in the baseline, which ultimately leads to a higher count of false positives. Other often-named challenges in the context of MCPS are limited sources of energy and constrained computational power. Often, especially in the case of wearable devices, those resources are already utilized by the device's primary purpose, so few resources remain for the intrusion detection algorithm. Moreover, even if one may argue that some wearables have an easily changeable battery, the need for energy-saving algorithms and protocols cannot get clearer for implantable medical devices (IMD). Consequently, motivated by these considerations, several research endeavors focus on energy and resource-efficient intrusion detection approaches (e.g., [24], [25], [26], [27]).

A third category, sometimes also considered a subcategory of anomaly-based IDSs, is specification-based intrusion detection [28]. In this approach, all possible behaviors of the given medical device are specified. The

device's operation is then monitored. An alarm is triggered if the device transitions to an unspecified operating state. We decided to follow Mitchell and Chen's definition and consider specification-based intrusion detection as a standalone category [29] because the medical sector offers unique possibilities for specifications. It, therefore, enjoys special attention in the field of MCPS intrusion detection. The researchers promise that it combines the advantages of signature-based and anomaly-based detection, namely the ability to identify previously unknown attacks while limiting false positives and requiring less computational power than ML-based anomaly detection. However, very detailed knowledge of the monitored medical device is needed, and this approach is therefore associated with a high initial implementation effort.

Furthermore, hybrid approaches combine two or more variants into a new approach. Here it is essential to state that several authors combined different ML algorithms and called their approach hybrid. Since the distinction to, e.g., ensemble learning methods was too small from our point of view, we did not follow this subsumption. Therefore, we only classified an approach as hybrid if it comprised variants from different main categories (e.g., anomaly-based and signature-based).

Where is the attack detection system located?

Classically, there are two locations where attack detection systems are usually placed. On the one hand, a host-based IDS (HIDS) runs on the device and monitors the station's operating system, processes, or logs. On the other hand, a network-based IDS (NIDS) inspects the network traffic and often monitors the traffic of all devices connected to the network. The locality of the NIDS, particularly in segmented networks, can, in turn, influence its effectiveness and therefore be decisive. Both locations have their advantages and disadvantages. An NIDS is able to detect external threats at an early stage, but the mass of data can cause limitations, especially in large networks. While an HIDS might not notice external threats as early as an NIDS, it might detect malicious insiders that remain hidden to NIDSs [28]. In addition to this distinction, we observed a third location often chosen by the researchers in the MCPS domain: cloud or cloudlet-based IDSs. Here, too, hybrid approaches are conceivable and in other sectors pervasive.

What kind of data is analyzed by the attack detection approach?

This question often interrelates with the location of the detection system (or at least with the collector's location). At the network level, detection approaches might use metadata of captured packets or analyze the whole packet, more or less understanding the entailed sector-specific protocols. At the host level, various information about the operating system, processes, or log files can be evaluated. Of course, all this data can also be conglomerated in a cloud to be processed centrally.

What attack scenario is the primary focus of the detection system?

Frequently, detection approaches specialize in the defense against specific scenarios. This is because an outside attack is detectable by different indicators than an insider abusing valid privileges. We identified the scenarios with the greatest research interest and those that may be underrepresented in current research.

Which datasets and sources are utilized to evaluate the effectiveness of the detection approach?

Publicly available datasets make the detection approaches of different researchers comparable. Sometimes, however, researchers cannot find a dataset that fits their use case and look for alternatives. Some build test environments with simulators or real devices, while others generate data in other ways. We examined the approaches and the most used datasets and -sources in the field of MCPS attack detection.

What approaches go beyond detection and also include preventive measures?

It is often of particular research interest to not only detect but also mitigate attacks as quickly as possible. This is also appealing in healthcare, as any attack might endanger human life. On the other hand, one of the biggest challenges in the field of attack detection, especially in the case of anomaly detection, is the false-positive rate. This gets even more relevant if automated mitigation measures are taken. By reviewing the relevant articles, we explored how researchers address the potentials and risks in this regard.

3.2 Identification of Research

To capture the current state of research, the variety of terms used in the literature for networked medical devices alone necessitated a structured approach. We were not the first to find that IT security terms and definitions diverge in healthcare. Athinaiou et al. surveyed the IT security language and observed that definitions of concepts differed in health environments [30]. We identified 22 terms used in reference to such systems (Connected Health, Connected Healthcare, Digital Healthcare, e-health network, Healthcare IoT-based Systems, Healthcare 4.0, (Industrial) Healthcare Systems, Internet of Health Things, Internet of Healthcare Things, Internet of Medical Things, IoMT, IoT-Health, Medical Cyber-Physical Systems, Medical CPS, MCPS, Medical Information Systems, Medical Internet of Things, Medical IoT, Medical Sensor Networks, Networked Healthcare, Networked Medical Devices, (Smart) Medical Devices).

While there are no precise definitions, it is our impression that term combinations of *medical/health* and *internet of things* (IoT) like IoMT, mIoT, or IoHT have been used to refer not only to medical devices in hospitals but also to devices used to monitor specific health values at home. In contrast, the term MCPS was used almost exclusively for medical devices in a hospital context.

Table 1: **Digital libraries consulted for study selection.**

Electronic Source	Results
ACM Digital Library	139
IEEE Xplore Digital Library	2116
ScienceDirect	933
Springer Link	1356
PubMed	810
Total	**5354**

However, this observation did not apply to all publications, and we noticed a convergence of the device classes. Researchers hypothesize that all sensors monitoring patients' health parameters in hospitals will be connected to local gateway devices in the future [31]. This evolution can already be observed and is the reason for the prevalence of so-called medical device gateways in hospitals that connect medical devices to the hospital network. According to ENISA, these devices presently account for 34% of all devices in the healthcare sector [5]. Besides, it is quite similar to the convergence of general IoT and cyber-physical systems (CPS). NIST established in a special publication in 2019 that the concepts of CPS and IoT have become more and more equal and that the definitions can recently often be used interchangeably [32]. However, to clarify that this work focuses on detecting attacks on medical devices available for hospitals, we used the term MCPS.

In addition, we identified five expressions describing the detection of attacks (Detection, Network Security Monitor, Network flow, IDS, and Intrusion Prevention). The combined search strings were employed to search five electronic libraries. The results per library can be seen in table 1. In total, we obtained 5354 papers matching our search strings.

Table 2: **Inclusion criteria used during study selection.**

#	Inclusion Criteria
1	Research is peer-reviewed
2	Study is published and available in full length
3	Research is within the focus area (intrusion detection in the context of MCPS)
4	Study has been published before March 2023

Table 3: **Exclusion criteria used during study selection.**

#	Exclusion Criteria
1	Study is not written in English
2	Study is a duplicate result
3	Paper has been retracted
4	Paper published other than conferences, journals, patents, technical reports
5	Study lies outside the IT security domain
6	Study focuses on medical devices without hospital context

3.3 Selection of Primary Studies

Following Kitchenham et al., two authors performed a two-step screening of all obtained papers and selected those relevant to the research topic. To make the process comprehensible and verifiable, we defined the selection criteria in tables 2 and 3. In the first quantitative screening, the title and abstract of the publications were evaluated. The vast majority of the papers was excluded in this step. For the qualitative screening, 358 papers remained. The high rejection rate is attributable to the fact that many intrusion detection synonyms are also used in medical regard. Two examples of major fields in medical research are disease detection and monitoring of patients' health parameters utilizing various medical devices. Unfortunately, those terms could not be excluded from our search terms for obvious reasons, which led to a high rate of false positive results.

In the following qualitative screening, the full-text versions of the 358 papers have been consulted to single out those relevant to our research questions. The papers were screened by two researchers independently, and the resulting selection of included papers differed. The agreement has been measured using the Cohen Kappa statistic [33]. The initial value of the Kappa statistics was 0,826. Afterward, all disagreements were discussed and resolved. In the end, 118 papers were selected for data extraction.

3.4 Data Extraction and Synthesis

For data extraction, the remaining studies were read in full and categorized by the research questions defined in section 3.1. Thereby, we were able to answer the questions as comprehensively as possible. Here we followed the recommendation of Kitchenham et al. and assigned one researcher as the data extractor and the other as the data checker. Emerging disagreements have been discussed, and all researchers have agreed on the final classification.

Finally, the results of the review were summarized. In the following section, we will provide the gained insights.

4 Results

We identified 118 papers that could contribute to answering the research questions. However, not every paper could be consulted to answer every research question. One example is the study by Ardito et al., who outline a framework but did not implement it or test it using a dataset [34]. Therefore, while we were able to use this publication to evaluate the proposed detection approach (RQ 1), it was not suitable for answering the question about the used data sources (RQ 5). The exact number of papers included in the evaluation of each research question is indicated in each subsection.

Table 4: **An overview of the different detection approaches.**

Detection Approach	Paper	Count
Anomaly-based	Ahmed et al. [35], Akram et al. [36], Akshay Kumaar et al. [37], Alamleh et al. [38], Almaiah et al. [39], Alotaibi [40], Alrashdi et al. [41], Ardito et al. [34], Arfaoui et al. [42], Ashraf et al. [43], Astillo et al. [44], Awotunde et al. [45], Ayoub et al. [46], Balasubramanyn et al. [47], Basharat et al. [48], Bassene and Gueye [49], Cai et al. [50], Carreon-Rascon and Rozenblit [51], Chowdhury et al. [52], Fernandez et al. [53], Ferrag et al. [54], Fouda et al. [55], Gao and Thamilarasu [56], Ghourabi [57], Gupta et al. [58], Gupta et al. [59], Gupta et al. [60], Hady et al. [61], Hajder et al. [62], Hameed et al. [63], Hameed et al. [64], Hameed et al. [65], Haque et al. [66], Haque et al. [67], He et al. [68], Hei et al. [69], Hussain et al. [70], Igbe et al. [71], Iqbal et al. [72], Kamble and Gawade [27], Karthick Kumar et al. [73], Khan et al. [74], Khan et al. [75], Kilincer et al. [76], Kintzlinger et al. [77], Kumar et al. [78], Kumar et al. [79], Kumar et al. [80], Li et al. [81], Liaqat et al. [9], Mahler et al. [82], Manimurugan et al. [83], Mishra and Bagade [84], Mohammed and Aiheeti [85], Mowla et al. [86], Muhammed et al. [87], Nandy et al. [31], Nayak et al. [88], Newaz et al. [89], Newaz et al. [90], Odesile and Thamilarasu [26], Otoum et al. [91], Otoum et al. [92], Otoum et al. [93], Panagoda et al. [94], Priya et al. [95], Radoglou-Grammatikis et al. [96] Radoglou-Grammatikis et al. [97], Rahmadika et al. [98], Ram and Kumar [99], Rao et al. [100], Rao et al. [101], Ravi et al. [102], Rehman et al. [103], Saba [104], Saheed and Arowolo [105], Said et al. [106], Salem and Mehaoua [107], Schneble and Thamilarasu [108], Sehatbakhsh et al. [109], Sharma et al. [110], Singh et al. [111], Siniosoglou et al. [112], Spegni et al. [113], Tabassum et al. [114], Tahir et al. [115], Thamilarasu et al. [116], Thamilarasu et al. [24], Thapa et al. [117], Toor et al. [118], Wa Umba et al. [119], Wagan et al. [120], Wahab et al. [121], Wang et al. [122], Yan et al. [123], Zaabar et al. [124], Zachos et al. [125], Zubair et al. [126]	98
Specification-based	Abdulhammed et al. [127], Choudhary et al. [128], Fang et al. [129], Mitchell and Chen [29], Mitchell and Chen [130], Li et al. [131], Raiyat Aliabadi et al. [132], You et al. [133], Zhang et al. [134]	9
Signature-based	Boujrad et al. [135], Meng et al. [25], Mpungu et al. [136], Zhang et al. [137]	4
Hybrid	Begli et al. [138], Chen et al. [139], Dupont et al. [140], Meng et al. [141], Kolokotronis et al. [142], Lakka et al. [143], Tariq et al. [144]	7

RQ 1 - Utilized detection approach and employed technology

We identified three main approaches in the context of MCPS attack detection: anomaly-based detection, signature-based detection, and specification-based detection.

As shown in table 4, most researchers focus on an anomaly-based detection approach (98). The majority proposes an ML algorithm they tweaked to be most suitable for MCPS (42). Often, the approaches consist of an optimized feature/dimensionality reduction algorithm and an ML algorithm that performs the actual detection. While most researchers substantiate why their approach works best (e.g., Saheed et al. with their swarm-based approach [105]), others focus on optimizing parts of their approach. E.g., Priya et al. measure the benefits of different dimensionality reduction approaches [95]. The detection algorithm then classifies the traffic, flow, or packet into malicious/benign (binary classification) or even categorizes it into specific attack groups. One example of the latter is the work of Mowla et al., which attempts to identify an attack and classify the attack type [86]. Astillo et al. focus on one specific MCPS: a diabetes management control system consisting of three separate components: a sensor that steadily measures a patient's glucose level (continuous glucose monitor (CGM)), an insulin pump, and a controller. Their detection approach first estimates the blood glucose level of the patient. Thereafter, estimated and actual values are compared and derived as features. Eventually, the classification module evaluates if the current event cycle is anomalous [44]. Khan et al. criticize that researchers have so far focused on optimizing accuracy and false alarm rate while no attention has been paid to interpreting the prediction model. Therefore, they use an explainable model that provides information about the features leading to the prediction. Their motivation is to help security personnel to react timely and in the right way to an alarm and to increase trust in their detection model. They explain this is especially necessary for the healthcare domain since there are too few security experts [74].

20 of the papers compare several anomaly-based approaches to one another and assess the advantages and disadvantages of the approaches in the context of MCPS. E.g., Newaz et al. developed HealthGuard, which utilizes four ML-based detection techniques (Artificial Neural Network, Decision Tree, Random Forest, k-Nearest

Neighbor) [89]. The researchers compare the algorithms in terms of accuracy, precision, recall, and F1-score (test accuracy considering precision and recall). 14 researchers combine different anomaly-based approaches to a new, amalgamated approach. While most state how their approach improves the anomaly detection, Kintzlinger et al. emphasize that their proposition of a combination of ML algorithms and statistical methods performs worse than the use of statistical methods alone [77].

Another repeatedly seized approach is Federated Learning (17) which researchers use to address the challenges of healthcare data privacy (e.g., Otoum et al. [92], Thapa et al. [117], Ferrag et al. [54]). It is a machine learning technique that has recently attracted much attention – not only in medical applications – because it protects data privacy. Other ML approaches often store data centrally without taking privacy-preserving measures. This turns these central data stores themselves into lucrative targets. In contrast, Federated Learning establishes a global learning platform that combines the knowledge of locally available models. The process of training an algorithm runs over separate decentralized models. Local datasets are used without revealing private data. Federated learning can thus preserve the training dataset on the devices so that the patient's data is not needed for training on the server side [93]. While several researchers include one network segment or a whole hospital in a local model, Gupta et al. propose a digital twin for each patient and train their local model on it. The advantage is that all collected data belong to a single patient, and the researchers can correlate more parameters [59].

Specification-based detection approaches are the second largest group, though by a large margin (9). We present three examples in the following: To detect maliciously acting devices instead of attacks, Mitchell and Chen devise a behavior specification-based approach. They define behavior rules and derive attack states from there. Subsequently, the researchers develop state machines. The authors promise this approach could detect unknown attacks while keeping the overhead and false positive rate low [130]. Refining this work, Abdulhammed et al. create a hardware approach (Field Programmable Gate Array (FPGA) chip) that employs behavior rules to detect anomalies [127]. Their approaches address the resource constraints of MCPS. Fang et al. also observed and analyzed the behavior of the monitored devices. They suggest a combined approach of fuzzy core vector machine and rough set (RS) as preprocessor (peculiarity: RS acts as a filter for apparent abnormal behavior) [129].

Exclusively signature-based approaches propose only four researchers in their publications. Meng et al. and Zhang et al. are two examples: Meng et al. cover the topic of decentralized detection for privacy reasons and outline a decentralized signature-based detection approach [25]. Zhang et al. combine the open source IDS *Snort* and the vulnerability scanner *Nessus* for an attack intention prediction [137].

Although signature-based detection may seem to be the least pursued approach, some researchers include it in a more general security strategy, combine it with another method (hybrid), or use it as a means of comparison. Dupont et al. wrote a protocol dissector for the IDS Forescout SilentDefense [140]. Magomedov recommends a signature-based approach for identified DICOM vulnerabilities [6]. Nguyen et al. designed a secure logger for medical devices with some detection capabilities. It consists of a dongle attached to the medical device that sends data to a remote cloud. The detection component focuses on packet or sequence tampering. Contrarily, the researchers consider compromised medical devices or devices sending compromised logs out of scope [145]. Radoglou-Grammatikis et al. compare their ML-based approach to Suricata loaded with attack signatures of Cisco Talos for the IEC 60 870-5-104 protocol. The signature-based approach performs better than most anomaly-based solutions presented in their work [97].

RQ 2 - Data collection and processing Locations

Researchers choose different locations and thus varying data sources for their IDSs. We differentiated between the device, network, and cloud/cloudlet locations and combinations of two out of those (figure 1). It is essential to state that we chose the location network if network traffic was seized, the location device if data was collected and processed on the device (e.g., log data), and the cloud if data was collected in or from cloud services.

Most researchers select the network as the sole location for their approach (52%). While a majority chooses a classical IP-based NIDS approach, some utilize particular circumstances of the healthcare sector or a specific MCPS. For instance, Gao and Thamilarasu propose a gateway device for an IMD and its programmer device. It acts as a man-in-the-middle and is supposed to detect attacks between those devices [56]. Mahler et al. developed an IDS specifically for a CT device. It intercepts traffic between the host pc and the device (on the can bus) [82].

The location cloud(-let) was chosen by eleven percent of researchers. This was often the case if researchers gathered health data from a manufacturer's cloud. Examples are Gupta et al. [59] and Newaz et al. [89], who correlate different vital signs of patients. The approach of Gupta et al. stood out since they took the first steps in matching network data and health data. From this, they assess the monitored user's behavior to detect abnormalities [58].

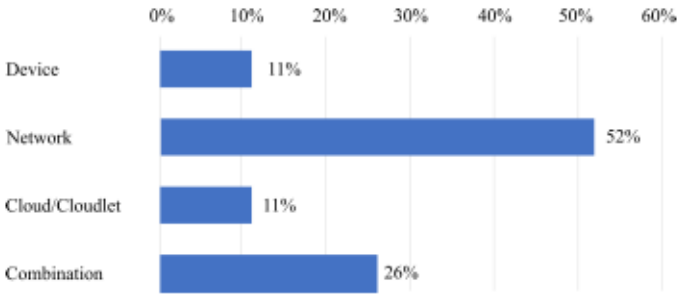

Figure 1: **The chosen data collection location for the different IDS approaches.** In total, 117 papers were analyzed. "Combination" consists of a mixture of two locations, where device and network make up 12%, device and cloud 4%, and network and cloud 10%.

Similarly, eleven percent of researchers opted for a pure on-device approach. To meet the special requirements of the healthcare sector, many researchers focus on lower resource constraints while preserving the patients' privacy. A particular advantage of the location device is that researchers can benefit from the special conditions of hospitals and clinics. E.g., Ardito et al. tailor their approach to an electrocardiography (ECG) device and use its user interface to display warnings in case of an anomaly. Then, feedback is requested from the treating physician. In this way, the physician is warned as quickly as possible in dangerous situations, and in the event of a false positive result, the effects can be limited [34]. A disadvantage of the location is that to implement a HIDS on a medical device, most researchers rely on device-specific knowledge or access to source code. This limits the transferability and scalability of the approach in many cases

Adding combinations of the locations device & network (12%) and device & cloud (4%), the number of researchers who combine the on-device approach with another location (16%) is higher than the number of researchers who use a pure on-device approach (11%). Thereby, in total, 27% of researchers decided to include device-specific information in their detection approach. Astillo et al. present one example of a combination of device and cloud. In their federated learning approach, they collect the data directly from the Continuous Glucose Monitor and process it on the controller unit of the MCPS. To share the knowledge between similar setups (other diabetes management control systems), only a submodel is generated on the device and subsequently fused to a central model on a cloud server [44]. The approach of Mitchell and Chen is a combination of host-based and network-based detection. Every node in the network acquires a set of behavior rules and can monitor the behavior of its trusted peers. So every medical device is monitored while it is also part of the detection approach [130]. Meng et al. suggest to perform the detection on every node individually and recommend a blockchain as an exchange platform for necessary signatures and a list of blocked nodes. As every node could add signatures to the chain in this scenario, the authors propose a centralized trust management scheme [25].

Besides the aforementioned categories, we could also observe that some researchers neglect the location choice of data collection. Instead, they base their detection approach on existing datasets (further elaboration in section 4.5) in computing platforms and simulation environments such as Matlab and Simulink. In this case, the dataset dictates the collection location. Examples are: Akram et al. [36] and Begli et al. [138]. Others combine the toolboxes with different simulators or platforms. Chen et al. employ Matlab and a cloudlet mesh simulator to calculate and evaluate the optimal number of collaborating IDSs in their cloudlet mesh approach [139]. In contrast, other researchers embed the proposed detection approach in a holistic security concept for a realistic hospital environment and even consider hospital network specifics.

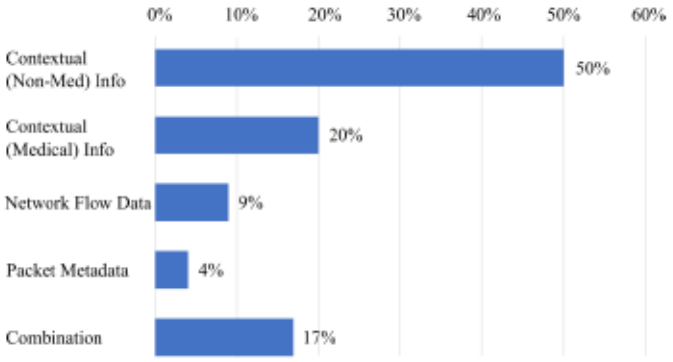

Figure 2: **The type of data from which the features and characteristics were obtained.** In total, 114 papers were applicable for the evaluation. "Combination" consists of a mixture of two data types, where network flow data and contextual (non-medical) data make up 13%, metadata and contextual information 3%, and metadata and network flow data 1%.

One example is Lakka et al., who describe an incident management approach, complementing their swarm-based detection with signature-based detection and consolidating the data in a hospital SIEM. A layered model outlines what information is collected where, sent where, and processed where [143]. Khan et al. also consider how their approach could be rapidly deployed in many hospitals. To this end, they have developed a framework for deploying their approach as Infrastructure as a Service in the cloud and as Software as a Service in a hospital network [75].

RQ 3 - Included features and characteristics of the leveraged data

In contrast to the IDS locations, we observed a higher variety in the examined features (figure 2). The majority of researchers base the detection on non-medical contextual information (50%), i.e., analyze technical data and transfer gained IDS-insights from other sectors to the medical sector. One often-used approach is the analysis of the network packet's contents. The medical sector is particularly interesting for ML-based detection approaches because of the resource constraints and the data masses generated by MCPS. Therefore, often observed research focuses are feature selection and hyperparameter tuning (e.g., Akshay et al. [37], Schneble and Thamilarasu [108]).

Remarkable is the high count of papers using contextual medical information from which researchers derive indicators of an intrusion. 20% of papers base their detection approach solely on such information. Mitchell and Chen were the first to incorporate medical context and correlations for attack detection (e.g., one proposed rule for conspicuous behavior is if the pulse is above a certain threshold during an analgesic request of the patient) [29]. Siniosoglu et al. utilize medical data such as ECG and arterial blood pressure [44]. Newaz et al. relate health values from different devices and interpret the results. They hypothesize that an attack usually targets one device at a time and that a deteriorating state of health should simultaneously affect various measured health values. If only single values deviate, they infer that

this data must have been manipulated. E.g., if the patient's oxygen level drops due to health reasons, her heart rate would naturally also decrease. So if only one of the values changes, the IDS will detect an anomaly and raise the alarm [59]. Hady et al. propose a packet comprehension functionality: Their models recognize the heart rate, respiration rate, systolic blood pressure, diastolic blood pressure, blood oxygen, and more from captured network traffic [61].

Others utilize network flow data (9%) or packet metadata (4%) and claim that this is more suitable than inspecting all packets. Besides the already mentioned data masses in the health sector, some researchers give additional reasons. Fernandez et al. argue, for example, that their primary motivation for using network flows is the more and more encrypted data sent over the network, due to which inspecting packets would be pointless [53].

Several researchers combine two types of features in their detection approach to identify attacks (17%). One example is the expansion from the field of disease classification to attack detection on medical data, as Haque et al. pledge their approach can do both [66]. Siniosoglou et al. leverage this approach and propose two supplementary models: one model to detect intrusions from network flow data and one model to detect anomalies from healthcare data [112]. Sehatbaksh et al. and Rao et al. follow entirely different approaches. The former propose to use the electromagnetic (EM) signals generated by the monitored medical devices during operation to distinguish between normal and malware-infected MCPS [109]. The latter suggest monitoring system operations. They hypothesize that a malware infection is identifiable by monitoring processes and other system parameters (especially the execution time) of a medical device [100].

RQ 4 - Adversarial focus

Many researchers limit the applicability of their work by making assumptions about attack types, targets, and locations, among other things. We have investigated which defense scenarios the detection approaches focus on. Here, we differentiated between external threat actors, malicious insiders, attack scenarios utilizing malware, and approaches focusing on detecting more than one attack scenario (figure 3).

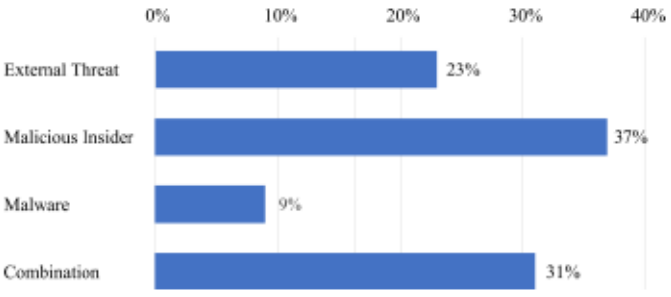

Figure 3: **The detection approaches focused on different attack scenarios.** In total, 113 papers were analyzed. Combinations consist of approaches promising to detect external and internal threats (25%), external and malware threats (3%), and internal and malware threats (3%).

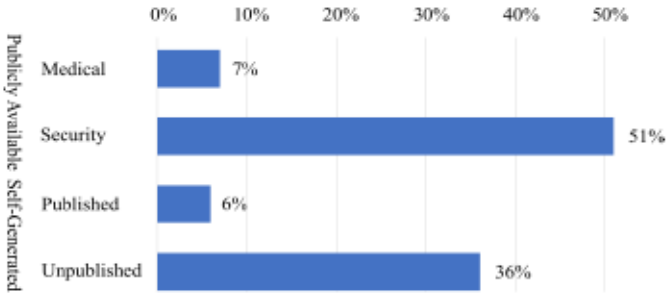

Figure 4: **The four categories of datasets and -sources encountered in the reviewed publications.** In total, 108 papers were analyzed.

Most researchers concentrate on insider scenarios (37%) or the combination of internal as well as external threats (25%). 23% focused on the sole detection of external threats. 9% of the researchers centralize the detection of malware infections. It is essential to state that we also included such papers in the category of insiders that do not explicitly mention such a specific attack scenario as a limitation, but require an attacker to have access to the network or a device (e.g., the attacker is able to spoof a mac address or the drug dosage is monitored for manipulations). So an external attacker that has already compromised an MCPS and can be detected as late as he laterally moves in the network or interferes with the normal function of the MCPS, is considered an insider in our classification scheme. The behavior-based approach of Fang et al. contrasts this scheme, as it promises to defend against external attackers. The model that they call *detecting illegal behavior (DIB)* focuses on the detection of maliciously acting accounts and devices (e.g., accounts that have been taken over through shoulder-surfing attacks) [129]. As it is technically impossible to differentiate between such a compromised account and a real insider sending malicious commands, we decided to follow our definition. While most other behavior-based detection approaches concentrate on the detection of insiders, Mitchell and Chen additionally claim to be able to detect malware, as they estimate malware to change the behavior of an infected device as well [130].

RQ 5 - Datasets used for validation

Researchers need data to train and test detection approaches (especially if they are ML-based). There is an additional benefit when multiple researchers use the same dataset, as the different detection methods become comparable. Choosing a dataset fitting the task is crucial since datasets are generated for specific purposes. We have organized the used datasets and -sources into two categories, each with two subcategories, as shown in figure 4. The first category includes approaches that utilize publicly available datasets. Here we differentiate between security and medical datasets. The second category deals with publications that have created their data themselves. While datasets from some approaches are available to the community, many remain unpublished.

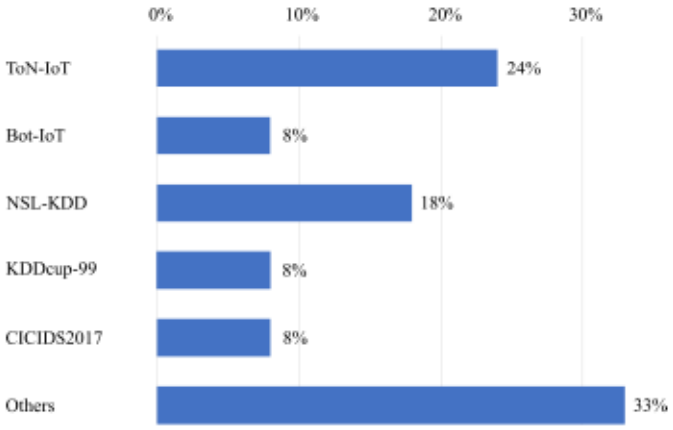

Figure 5: **The different publicly available network traffic datasets used.** During our analysis, we encountered 18 different datasets that were used 72 times. Some research groups used more than one dataset and were therefore assigned to more than one category.

Publicly Available Datasets

In non-health domains, researchers usually train and test novel IDSs utilizing publicly available security datasets. Our analysis shows that 51% of the researchers also follow this approach. Figure 5 presents the different datasets. 8% of the researchers in this group utilize the KDDcup-99 dataset. This dataset was developed for the KDD-cup competition in the year 1999, whose goal was to develop an NIDS [146]. Several problems, such as redundant records, have been reported, and the successor, the NSL-KDD dataset, was released in 2009 [147]. 18% of the researchers in this group use this dataset. Both datasets contain several IT protocols such as HTTP, SMTP, and FTP. Among others, Khan et al. bemoan the deficiencies of missing the latest attack vectors in the NSL-KDD dataset [74]. The Canadian Institute for Cybersecurity (CIC) published several datasets promising to have more recent attacks resembling real-world data. Their top priority varies from dataset to dataset. In the 2017 data set, they provided realistic background traffic and simulated the behavior of 25 users [148]. In the 2018 dataset, they focused on insider attacks and provided system logs of every machine [149]. Most researchers relying on CIC datasets in the examined works use the CIC IDS 2017 dataset (8%). It consists of raw packets in PCAP files as well as labeled flows [148]. However, none of the presented datasets comprise IoT or MCPS protocols. The University of New South Wales (UNSW) fill the gap with their datasets ToN-IoT [150] and Bot-IoT [151]. In these datasets, a real-world network environment containing both IT and IoT devices was mimicked. Besides the network traffic, the researchers from UNSW provide Windows and Linux audit traces and telemetry data for the IoT services. Their IoT testbed consisted of various devices, among others: a fridge, a garage door opener, a thermostat, a GPS tracker, a motion sensor, and a weather station [150]. MCPS, however, have not been included. The ToN-IoT dataset, by 24%, is the most used dataset, while the Bot-IoT dataset is employed by 8% of the researchers that rely on publicly available datasets.

Some researchers employ the medical sector only as motivation and ignore the discrepancy between real network traffic in hospitals and the datasets they choose to support their detection approach. Others point to the absence of MCPS traffic in the dataset and handle the inadequacy differently. For instance, Schneble and Thamilarasu explain that in the context of MCPS, two aspects are crucial to keeping the detection latency low: Feature selection and reducing the amount of data processed by the IDS. Hence, to test the effectiveness of their feature ranking and selection algorithm, they consult the MNIST digit recognition dataset. This dataset contains 60,000 handwritten digits. They choose this dataset, among other reasons, because of the large feature space and the easy access to the data [108]. In contrast, Ferrag et al. explicitly determine the MNIST dataset unsuitable for training and testing IDSs in the context of medical devices [13]. Hameed et al. state that their approach is only applicable for detecting MCPS in a real environment if it is properly adapted prior to deployment [63]. Tabassum et al. use the datasets KDDcup-99 and NSL-KDD and merge their self-generated IoT traffic to cope with the missing IoT traffic in named datasets. However, they do not explain which MCPS they employed for the generation [114].

Some researchers harness medical datasets containing patient and medical data to identify attacks from those datasets (7%). The most commonly used dataset is the MIMIC III dataset, which contains health-related data of forty thousand patients who received intensive care in a hospital in Israel [152]. 25% of the researchers in this group used this dataset. While the researchers found the specifics of the medical data particularly valuable for attack detection, the drawback is that none of these freely available medical datasets contain attacks. Therefore, alternative ways must be found here as well. One idea given by Siniosoglou et al. is to use two distinct datasets to train their neural network: A publicly available medical dataset, and the UNSW-NB intrusion detection dataset for network flow data [112].

Self-Generated Data(sets)

Many researchers generate their own data(-sets) and work with that data without publishing it afterward (36%). While this results in the fact that subsequent studies cannot be compared to their work, the reasons given are manifold. On the one hand, this data often results from cooperation with hospitals and could reveal real patient data. One example is Boddy et al., who captured network traffic in a UK hospital and depict the complexity of the network infrastructure in a visualization approach [153]. Even if this data could be anonymized, many argue that hospitals prefer to be on the safe side and not risk the exposure of any patient data. On the other hand, the data might be especially suited to an approach or just randomly generated, as in the work of Mitchell and Chen. They generated random data following their devised state machine [130]. This data would have had no benefit for any other researcher, as their states are unique to their approach.

As we already addressed in section 4.2, researchers used computing and simulation environments to test their new attack detection algorithm on existing datasets. An-

other approach is to utilize simulators and frameworks to model an even more realistic MCPS environment. Among these approaches are those designed for a medical environment and those whose original purpose is different. Two examples of non-medical simulators are presented in the following. Meng et al. operate a publicly available tool to generate attacks on wireless networks. The attacks are not specifically adapted to MCPS environments [25]. Thamilarasu et al. employ Castalia, a simulator for WSN and WBANs, in several papers [56] [26] [116]. Such toolboxes and frameworks originally developed for other purposes have certain limitations regarding MCPS simulation. Therefore, several researchers adapt various open-source medical device simulators to their needs or implement their own medical device simulators. Astillo et al. operate the UVA/Padova Type 1 diabetes simulator that has been approved by the U.S. Food and Drug Administration (FDA) in their testbed and generated their test data with it. They also use an extended simulator version to induce artificial attacks [44]. Sehatbakhsh et al. leverage open-source code to deploy a syringe pump on various architectures. They found a buffer-overflow vulnerability in the syringe pump's source code that they were able to exploit [109]. Raiyat Aliabadi et al. employ OpenAPS, an open source Smart Artificial Pancreas [132]. They use fault injection as the source for unknown attacks.

Recent research efforts concentrate on the standardization of medical device inter-connectivity to address the heterogeneity of network protocols used by medical devices in hospitals mentioned in section 1 This is not only an IT security challenge. One project that has already made some progress is the community implementation of an integrated clinical environment, OpenICE. It provides a framework for the integration of medical devices into an integrated clinical environment (ICE). The developers even promise to be able to connect legacy devices to their ecosystem. For that, they developed adapters for those devices and a novel network protocol [154]. Some security researchers propose IDSs for networks based on OpenICE. Li et al. use OpenICE to simulate future medical devices and accomplish a data flow analysis in an OpenICE network [131]. Fernandez et al. analyze network flows of malware outbreaks in such environments [53].

To mimic real-world hospital conditions even better, many researchers employ actual medical devices in a testbed. Figure 6 shows the different devices. While the medical devices most used are blood pressure sensors (12%) a clear favorite could not be determined. Various research groups cover multiple devices. A protruding example is the testbed of Fang et al. which contains 21 different medical devices and a malicious access point to capture network traffic. From the device behavior, they derive 21 behavior rules. Instead of attacking the devices, they define operation rules for each device and specify some operation rules as normal and the remaining as abnormal behavior [129]. This way, no real attacks are conducted. Instead, some behavior is defined as malicious. The detection system of Kintzlinger et al. is explicitly designed for attacks directed at programmer devices for implantable cardioverter defibrillators. They cooperated with two cardiology experts from a university medical center to create malicious programmings [77].

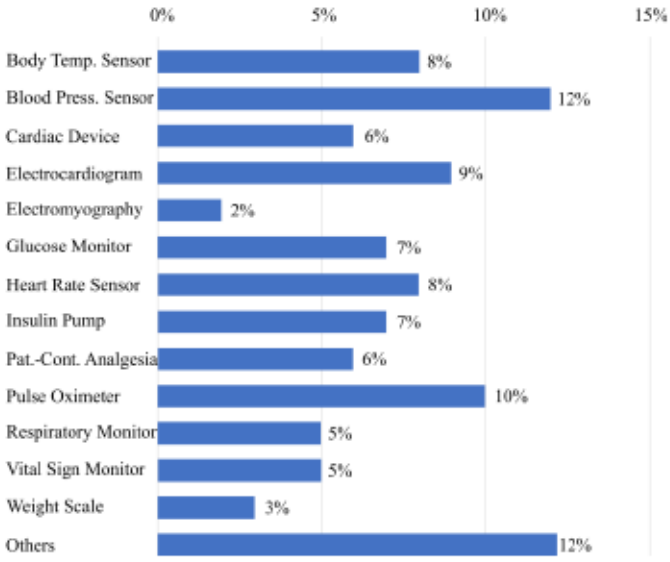

Figure 6: **The different medical devices deployed in the testbeds of researchers.** A total of 23 devices were used in 86 cases. Since several research groups analyzed more than one device, the percentages refer to the number of cases used (86) and not to the count of individual devices (23).

Yan et al. analyzed a medical shoe with 99 sensors attached. It is designed to detect the instability and balance of patients. The researchers statistically correlate the data of the different sensors in a shoe and, thereupon, identify attacks using anomaly detection [123].

Similar to Fang et al. before, we observed that many research endeavors were conducted utilizing household IoT devices rather than medical devices for data collection [129]. One example is Gupta et al., who built a conjoined testbed consisting of medical devices such as pulse oximeters and smart home devices like a fridge and a door sensor [58].

The same phenomenon occurs in the field of malware detection. Since there is little to no research on MCPS-specific malware, those researchers investigating malware outbreaks in clinical environments fall back on existing malware samples. Some utilize IT malware (e.g. Chowdhury et al. [52]), others use malicious android APK packages and explain that there are many mobile devices using android in hospitals [40].

Only six percent of research groups publish their generated datasets of MCPS-specific traffic. Nguyen-An et al. create an IoT traffic generator named IoTTGen. Their focus is smart home IoT as well as biomedical IoT. They analyze the behavior of smart homes and medical devices to build templates for those devices. The generator also allows adding new devices, if the traffic patterns are known. To generate anomalous packets, they extract attack traffic traces from the Bot-IoT dataset and inject them into their generated data. Since the Bot-IoT dataset does not contain MCPS-specific attacks, this generator cannot generate such attacks either. And since one finding in this work is that significant differences between the traffic of smart home devices and medical devices exist [155], it stands to reason that traces of attacks will differ as well.

Three dataset developers have focused on specific

protocols or technologies. Radoglou-Grammatikis et al. present a IEC 60 870-5-104 protocol dataset. It contains protocol-specific attacks. They use the IEC Testserver to deploy their MCPS devices without specifying what devices were modeled in detail [97]. Zubair et al. provide a Bluetooth-enabled medical device dataset. They record Bluetooth network traffic and generate flow data from such devices - explicitly excluding the collection of patients' exact health data. The attacks performed during the data recording are also Bluetooth-protocol-specific [126]. Hussain et al. focused on Message Queuing Telemetry Transport (MQTT) traffic and developed a tool for generating MQTT-based MCPS Traffic. The result is provided as an open-source dataset [70].

Two self-generated and published MCPS datasets have already been reused by other researchers: Ahmed et al. operate the Libelium Mysignals healthcare kit to generate their dataset. This kit provides a platform for the development of medical devices and eHealth applications. The researchers use three available health sensors in their testbed. Just their attacks are not medical protocol specific. Their ECU-IoHT dataset provides the recorded network packets in PCAP format, and network flows recorded using Argus [35].

Hady et al. record their testbed's network traffic, consisting of several small medical devices, and extract traffic flows and patient data from it. This recording is published under the name WUSTL-EHMS dataset. Among their carried-out attacks is the manipulation of medical data. Thereby, they directly integrated attacks on medical network traffic in their dataset, even if those are limited to spoofing and data modification from a MITM position. [61]. In addition, their WUSTL-EHMS dataset is the dataset from the category of self-generated and published datasets in our survey that has been used most often by other researchers to detect attacks on MCPS (4).

RQ 6 - Attack Prevention Endeavours

Attack detection is often closely associated with attack prevention since attackers and malware act fast, and a manual response often results in data loss or, in the case of the medical sector, patient harm. Therefore, a timely reaction is an often invoked point. Troublesome is that an automatic reaction based on a false positive might also harm a patient. One example is a higher-than-usual drug dosage given to a patient because of a life-threatening condition. If an IDS identifies this as an overdose attack and preventively interrupts medication delivery, the patient might die. Researchers have to consider these exceptional circumstances and come up with sector-specific solutions. Out of the 118 primary studies we reviewed, only 14 studies address attack prevention or mitigation. Others, such as Kumar et al, see the need for attack mitigation but choose to merely alert an administrator if an attack is detected and take no further action to thwart it [78].

Most preventive approaches (9) leverage software-defined networking (SDN). Fernández et al. propose a decision and reaction module in their approach. It consists of a rule-based decision component and a reaction and notification component. Utilizing network function visualization and SDN, medical devices can be isolated automatically. They emphasize that their approach is not to prevent the actual attack but to reduce the reach of the attack by preventing the attacker or malware from accessing more devices [53]. This strategy is also chosen by most other SDN-based prevention and mitigation approaches.

Others concentrate on preventing a specific attack vector in a specific environment. E.g., Thapa et al. utilize mitigation SDN rules to react to ransomware spread in an ICE [117]. Similarly, Bassene and Gueye focus their work on detecting DDoS attacks against hospital networks. They as well propose the utilization of SDN [49], but unlike the previous, their approach excludes entire subnets to counter this specific attack type.

Radoglou-Grammatikis et al. are part of the few who discuss the potential consequences of automatic preventive measures. They draw attention to the fact that an attack might come from a device still used for legitimate health-related operations. Their notification and response module weighs this risk of causing higher costs for the healthcare organization against the threat of the detected attacks. Eventually, it decides whether to isolate the device via automatically generated and applied firewall rules, or simply to report it to an administrator and ultimately have that administrator make the mitigation decision [97].

One example of a non-SDN-based approach is MedMon. MedMon is placed in a man-in-the-middle position between a controller and a wirelessly-connected medical device. Attack prevention works by jamming the identified malicious wireless connection. Regarding the option of a false positive, MedMon can be operated in different modes. If a valid connection from the controller to the insulin pump in the example of the researchers is jammed, the patient can manually deactivate the jamming [134]. Since insulin delivery does not have to occur within seconds, a patient can usually react to a warning. Therefore, this approach might be suitable for this particular device type. Contrarily, it might not be suitable for medical devices with other preliminary requirements.

Another non-SDN-based, innovative approach by Rao et al. proposes a new resilient design for MCPS. They suggest to employ different operating modes in the MCPS architecture. Threat mitigation is realized by automatically changing the operating mode based on calculated risk values. In 2018, they presented the idea of a so-called multi-modal system in the form of a pacemaker [100], while in 2019, they expanded the idea to an insulin pump [101]. Carreon-Rascon refined it and added self-healing capabilities to the system. In addition to the operational modes and threat mitigation policies, self-healing policies are proposed and linked to the tasks of the different modes. Once the steps of the active mode have been executed, it is possible for the MCPS to switch to the next lower-risk mode and execute the self-healing tasks coupled to this mode [51].

5 Discussion and Implications

Overall, many good reasons and motivations for intrusion detection in the medical sector have been published. It is clear that the healthcare sector is receiving special attention, and many argue that particular challenges require special solutions. In the following, we use the knowledge gained on the state of research to discuss obstacles and limitations of attack detection for MCPS. By deriving five future research topics (A-E), we would like to support prospective research projects to take a targeted direction.

5.1 Capturing the health sector-specific requirements for attack detection

The results show that the majority of researchers sees the difficulties of attack detection in the healthcare domain as detecting attacks with a low false positive rate, with as little computing power as possible while preserving patient privacy. The, by far, most popular approach is anomaly detection, while only a few researchers discuss options, how understaffed IT security personnel could handle results containing false positives. Overall, we observed that many researchers assume different conditions and circumstances in MCPS environments. Especially, technical differences between healthcare networks and those in other sectors have received little attention in research to date. Therefore, we identify the need to determine the requirements to which detection approaches in the medical field must adapt.

One main difference is the use of medical device gateways, which connect medical devices to hospital networks. Such a setup could lead to difficulties since, for example, agent-based approaches could not be easily deployed to such architectures. This also applies to innovative approaches that operate by combining local and global predicates. If, for instance, in an intelligent agent-based approach, such an agent can investigate a suspiciously-acting device, how could legacy devices be included without involving all their manufacturers? Furthermore, it is unclear to the research community if medical devices use encryption. As illustrated in section 4.3, some researchers assume that much of the traffic in a hospital is encrypted and propose a flow-based approach in response to this assumption. In contrast, other researchers like Hady et al. derive patient data from captured network traffic and assume that this traffic is sent unencrypted from medical devices to servers [61]. However, performing deep packet inspection (DPI) in the case of encrypted traffic (e.g., by implementing application layer gateways to break up encryption) is aggravated in the MCPS environment, as security engineers cannot easily place certificates on medical devices. Approaches that rely on DPI must take that into account. As a first future research topic, we see the need for a comprehensive study of the unique constraints, technical characteristics, and challenges of hospital networks, along with connected medical devices, in contrast to other IT and OT environments.

5.2 Creating MCPS-specific attack detection datasets

Researchers deal with the scarce information situation differently. Some leverage the diversity typical for the healthcare sector solely for motivational reasons. Others do not place much emphasis on where and how data is collected. Instead, they use existing, often outdated datasets that do not fit the field or their motivation and ignore any differences. Again other researchers find 'creative' ways to replicate individual features of hospital networks and test specific parts of their approach. One example is the debatable use of the MNIST digit-recognition dataset to reenact the high feature dimensionality in the health sector. Either way, detection based on real health-specific protocols is rarely conducted, leading to limited portability of detection approaches from outside the medical domain. Moreover, it leads to uncertainty about whether the supposedly most suitable approach will be the best fit in a real hospital environment. ML-based IDSs must, most certainly, be retrained to prevent an increased false positive rate in a real-world environment. Furthermore, even the more recent datasets (e.g., CIC IDS 2017/2018) are often unsuitable for detecting attacks on medical devices since they do not contain IoT or IIoT traffic. Even if such datasets exist (e.g., TON-IoT or Bot-IoT), they might not be suitable for the health domain, as we saw that MCPS traffic significantly differs from other IoT traffic. There are several health-specific protocols (e.g., HL7 and DICOM [6] [140]), but only exceedingly few of these protocols are part of the datasets examined. Problematically, current adversaries are increasingly attacking application-layer protocols, as discussed by Hussain et al. [70]. Another factor not covered in current datasets is that different real-world attackers would behave differently. While several researchers, among other things, focus on detecting port scans or denial-of-service attacks, APT attackers would act much more stealthy. The first steps of recognizing the differences in the attacker's modus operandi were taken by Mitchell and Chen, who consider this fact with their attacker archetypes (reckless, opportunistic, and random attacker) [130]. Thamilarasu, too, takes the attacker's behavior's impact on the effectiveness of the detection into consideration [116]. However, these researchers use simulations for their distinctive environment. A dataset containing such characteristics has yet to be developed.

In addition to the uncertainty about the transferability to the real world, the lack of fitting datasets limits the comparability of the approaches. When comparing ML-based approaches, it is common to compare performance indicators such as accuracy, precision, and recall. Many of the papers have calculated and reported the corresponding values in their evaluation. However, we have refrained from correlating papers based on these metrics in this paper. On the one hand, this is because often, not the same datasets were used for training and testing of the individual algorithms so that, at most, a small group of algorithms could be compared to each other. On the other hand, often further assumptions were made about attackers, attacks used, or the granularity of the classification (as described in section 4.4). These assump-

tions limit the applicability of the algorithms to single-use cases and further reduce the comparability to other algorithms. Sharma et al. reacted to this and built a modular framework with a benchmarking suite. This could help future researchers to easily test their new detection algorithms and compare them directly to the work of other researchers [110]. But since this framework represents a novelty, the community must first accept it. Furthermore, this suite also relies on the existing datasets, and while it takes an important step for comparability, it is not a comprehensive solution to this concern.

We also observed a trend to utilize distributed and federated learning approaches. It is crucial to point out that these models have specific requirements for datasets and that current datasets do not fulfill them.

In conclusion, the lack of appropriate datasets is a major obstacle to developing attack detection in the healthcare sector. From our point of view, an attack detection dataset should include three things to be suitable for the health domain. It should: (a) incorporate health-specific protocols, (b) model different attacker behavior, and (c) be suitable for specific scenarios and techniques, such as distributed and federated learning. The generation of such MCPS-specific datasets has to be addressed in the future. Therefore, we proclaim it as the second future research topic.

5.3 Advancing standardization projects

In addition to capturing the current state and the characteristics of the healthcare sector and mapping that state into datasets, we see the need to get to the root of the increasing diversity and the individual technology that makes IT security in healthcare so difficult. Therefore, efforts to standardize the sector's digital infrastructure must be intensified. Initiatives such as OpenICE are commendable. However, OpenICE was developed without consideration of security [145]. Additionally, intrusion detection in OpenICE-based networks will not be easily transferable to real hospital networks, as the network protocols are unique to the OpenICE environment.

The urgent need for standardization also applies to the device level. While it seems unsurprising to find the majority of researchers choosing the network as the data collection location (corresponding to the insights into the heterogeneity of medical devices discussed in section 1), many researchers are examining single medical devices and designing specially adopted host-based detection approaches. This suggests that despite the hurdles in this area, many researchers would like to take advantage of the insights that device data can provide. One example of the many possibilities is the implemented feedback reaction to a detected anomaly by Ardito et al. [34] discussed above. However, the heterogeneity of devices currently means that a separate solution is required for each type of device. That is why most researchers propose an HIDS focused on a single or few device types (e.g., a smart artificial pancreas [132], a smart-connected-pacemaker [100], or a diabetes management control system [44]). The transferability of these approaches to the real world and, in particular, scalability are problems that are often not addressed by researchers. Therefore, manufacturers should also incorporate security considerations into the development of devices. Device and log data has to be made accessible to security experts in a standardized way. We ascertain advancing standardization ventures, therefore, as the third future research topic. As explained, this applies to both the device level and the network level.

5.4 Connecting technical and medical data for attack detection

In addition to the technical aspects of healthcare networks, some researchers explored the potential of medical context in various forms for intrusion detection. The promised benefits were manifold. E.g., in the case of a syringe pump, medical context could provide insights into a too-high dose for a patient and thereby recognize not only technically novel attacks and malicious insiders but also simple mistakes of health personnel. However, any initial attacker efforts or intrusion attempts might go unnoticed in these approaches. An attacker is only discovered if a device is already compromised and she tries to manipulate the care process. We have presented initial approaches for combined detection based on network traffic and medical data. However, detection has taken place independently and based on unrelated data sources. The genuine and thorough integration of the medical context with the detection based on technical features and, thus, creating a holistic approach is the fourth future research topic.

5.5 Driving risk-aware attack prevention

During this survey, we observed that automatic intrusion prevention in the medical sector is an area that is handled even more carefully than in other sectors. The reason lies in the high stakes at risk: Lives depend on the system's proper functioning. While in other domains, a quick shutdown of a system that is most likely compromised may be just the right response in the risk assessment, an MCPS might still provide life-sustaining measures despite a compromise. Thus, the reaction must be weighed quite differently in this domain. As presented in section 4.6, there are very few research groups that address prevention and mitigation at all. The vast majority of them use the isolation capabilities that their SDN approach provides. While this does not necessarily mitigate the attack, it can stop potential lateral movement. Especially in the context of malware (esp. ransomware), this can be very valuable. However, the implications for the further functioning of isolated medical devices are rarely considered. Other endeavors propose individual solutions for single medical devices. While these preventive approaches are often innovative, they are tailored to the specific device type, and their risk assessment is not (easily) transferable to other devices. A plausible example is the IDS for an insulin pump, which notifies its user in case of an anomaly. The user can override the preventive measures and thus correct a false detection if necessary. This procedure would be fatal in the case of a pacemaker, for example, because here, it is important to react very promptly to anomalies. If the user is consulted first, it is questionable whether she can respond in a timely manner (or at all).

Since attackers and malware make no distinction between hospitals and other targets, these considerations and difficulties must not lead to a neglect of prevention. Just as in other fields, a quick but well-thought-out response in the event of an attack is essential in the medical field. Hence, a fundamental discussion about intrusion prevention in the medical domain, the sector-specific requirements, and how it can succeed despite the high risks has to be conducted. We conclude that this is the fifth future research topic.

6 Conclusion

Already in 2012, Clark and Fu denounced two challenges in the context of the security of medical devices: "(1) computer security researchers seldom have access to real medical devices for experimentation, and (2) the computer security community is largely disjoint from the biomedical engineering community." [156] These challenges persist to this day.

In this paper, we conducted a structured literature review by following the guidelines of Kitchenham et al. We found the synonyms for MCPS to be manifold and many of the security terms to be used in other respects in the medical domain. Most researchers focused on an anomaly-based detection approach at the network layer. The detection of malicious insiders was the primary focus. Several researchers used publicly available datasets for training and testing their algorithms. Others criticized the lack of suitable datasets and developed testbeds consisting of various medical devices. While some medical devices were used by multiple research groups, we observed no clear preference. Based on the results, we identified five research gaps. We discussed why it is necessary to examine the special conditions of hospital networks, the MCPS deployed within them, and the contrasts to other IT and OT environments. Furthermore, we see an urgent need for the creation of MCPS-specific datasets. Only with these sets researchers can attribute to the requirements and the unique possibilities of the healthcare domain. Alongside this, we see the need to support and expand MCPS standardization projects. Moreover, the medical domain offers an excellent opportunity to fortify attack detection based on technical features with medical context, thereby creating a holistic approach. Last but not least, a fundamental discussion should be held about the challenges of intrusion prevention in the medical domain and how it can succeed despite the high risks. We are confident that by countering these challenges, IT security in hospitals can be enhanced, and patients' lives can be protected.

References

[1] P. Bischoff, "Ransomware attacks on us healthcare organizations cost $20.8 bn in 2020," Comparitech. https://www.comparitech.com/blog/information-security/ransomware-attacks-hospitals-data/, 2021, [Online; accessed 25-October-2022].

[2] B. Filkins, "Health care cyberthreat report: Widespread compromises detected, compliance nightmare on horizon," *SANS Institute*, vol. 42, 2014.

[3] I. Gartner, "Gartner predicts by 2025 cyber attackers will have weaponized operational technology environments to successfully harm or kill humans," https://www.gartner.com/en/newsroom/press-releases/2021-07-21-gartner-predicts-by-2025-cyber-attackers-will-have-we, 2021, [Online; accessed 25-October-2022].

[4] L. Coventry, D. Branley-Bell, E. Sillence, S. Magalini, P. Mari, A. Magkanaraki, and K. Anastasopoulou, "Cyber-risk in healthcare: Exploring facilitators and barriers to secure behaviour," in *International Conference on Human-Computer Interaction*. Springer, 2020, pp. 105–122.

[5] E. U. A. for Cybersecurity, "Enisa threat landscape 2021—april 2020 to mid-july 2021," 2021.

[6] S. G. Magomedov, "Software for Analyzing Security for Healthcare Organizations," in *Futuristic Trends in Network and Communication Technologies*, ser. Communications in Computer and Information Science, P. K. Singh, G. Veselov, V. Vyatkin, A. Pljonkin, J. M. Dodero, and Y. Kumar, Eds. Singapore: Springer, 2021, pp. 181–189.

[7] D. E. Denning, "An intrusion-detection model," *IEEE Transactions on software engineering*, no. 2, pp. 222–232, 1987.

[8] I. Gartner, "Gartner market guide for operational technology security," https://www.forescout.com/gartner-market-guide-for-operational-technology-ot-cybersecurity/, 2022, [Online; accessed 08-December-2022].

[9] S. Liaqat, A. Akhunzada, F. S. Shaikh, A. Giannetsos, and M. A. Jan, "SDN orchestration to combat evolving cyber threats in Internet of Medical Things (IoMT)," *Computer Communications*, vol. 160, pp. 697–705, Jul. 2020. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0140366420312044

[10] M. Banerjee, J. Lee, and K.-K. R. Choo, "A blockchain future for internet of things security: a position paper," *Digital Communications and Networks*, vol. 4, no. 3, pp. 149–160, Aug. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352864817302900

[11] J.-P. A. Yaacoub, M. Noura, H. N. Noura, O. Salman, E. Yaacoub, R. Couturier, and A. Chehab, "Securing internet of medical things systems: Limitations, issues and recommendations," *Future Generation Computer Systems*, vol. 105, pp. 581–606, Apr. 2020. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0167739X19305680

[12] T. Tervoort, M. T. De Oliveira, W. Pieters, P. Van Gelder, S. D. Olabarriaga, and H. Marquering, "Solutions for Mitigating Cybersecurity Risks Caused by Legacy Software in Medical Devices: A Scoping Review," *IEEE Access*, vol. 8, pp. 84 352–84 361, 2020, conference Name: IEEE Access.

[13] M. A. Ferrag, L. Shu, and K.-K. R. Choo, "Fighting COVID-19 and Future Pandemics With the Internet of Things: Security and Privacy Perspectives," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 9, pp. 1477–1499, Sep. 2021, conference Name: IEEE/CAA Journal of Automatica Sinica.

[14] N. M. Thomasian and E. Y. Adashi, "Cybersecurity in the Internet of Medical Things," *Health Policy and Technology*, vol. 10, no. 3, p. 100549, Sep. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2211883721000721

[15] S. S. Hameed, W. H. Hassan, L. A. Latiff, and F. Ghabban, "A systematic review of security and privacy issues in the internet of medical things; the role of machine learning approaches," *PeerJ Computer Science*, vol. 7, p. e414, Mar. 2021, publisher: PeerJ Inc. [Online]. Available: https://peerj.com/articles/cs-414

[16] Y. Rbah, M. Mahfoudi, Y. Balboul, M. Fattah, S. Mazer, M. Elbekkali, and B. Bernoussi, "Machine Learning and Deep Learning Methods for Intrusion Detection Systems in IoMT: A survey," in *2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, Mar. 2022, pp. 1–9.

[17] F. Pelekoudas-Oikonomou, G. Zachos, M. Papaioannou, M. de Ree, J. C. Ribeiro, G. Mantas, and J. Rodriguez, "Blockchain-Based Security Mechanisms for IoMT Edge Networks in IoMT-Based Healthcare Monitoring Systems," *Sensors (Basel, Switzerland)*, vol. 22, no. 7, p. 2449, Mar. 2022.

[18] C. Eliash, I. Lazar, and N. Nissim, "SEC-C-U: The Security of Intensive Care Unit Medical Devices and Their Ecosystems," *IEEE Access*, vol. 8, pp. 64 193–64 224, 2020, conference Name: IEEE Access.

[19] M. Kintzlinger and N. Nissim, "Keep an eye on your personal belongings! The security of personal medical devices and their ecosystems," *Journal of Biomedical Informatics*, vol. 95, p. 103233, Jul. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1532046419301522

[20] P. Ghosal, D. Das, and I. Das, "Extensive Survey on Cloud-based IoT-Healthcare and Security using Machine Learning," in *2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*. Kolkata, India: IEEE, Nov. 2018, pp. 1–5. [Online]. Available: https://ieeexplore.ieee.org/document/8718717/

[21] S. M. Wa Umba, A. M. Abu-Mahfouz, T. Ramotsoela, and G. P. Hancke, "A Review of Artificial Intelligence Based Intrusion Detection for Software-Defined Wireless Sensor Networks," in *2019 IEEE 28th International Symposium on Industrial Electronics (ISIE)*, Jun. 2019, pp. 1277–1282, iSSN: 2163-5145.

[22] M. Wazid, A. K. Das, J. J. P. C. Rodrigues, S. Shetty, and Y. Park, "IoMT Malware Detection Approaches: Analysis and Research Challenges," *IEEE Access*, vol. 7, pp. 182 459–182 476, 2019, conference Name: IEEE Access.

[23] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," 2007.

[24] G. Thamilarasu and Z. Ma, "Autonomous mobile agent based intrusion detection framework in wireless body area networks," in *2015 IEEE 16th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, Jun. 2015, pp. 1–3.

[25] W. Meng, W. Li, and L. Zhu, "Enhancing Medical Smartphone Networks via Blockchain-Based Trust Management Against Insider Attacks," *IEEE Transactions on Engineering Management*, vol. 67, no. 4, pp. 1377–1386, Nov. 2020, conference Name: IEEE Transactions on Engineering Management.

[26] A. Odesile and G. Thamilarasu, "Distributed intrusion detection using mobile agents in wireless body area networks," in *2017 Seventh International Conference on Emerging Security Technologies (EST)*, Sep. 2017, pp. 144–149, iSSN: 2472-7601.

[27] P. Kamble and A. Gawade, "Digitalization of Healthcare with IoT and Cryptographic Encryption against DOS Attacks," in *2019 International Conference on contemporary Computing and Informatics (IC3I)*, Dec. 2019, pp. 69–73.

[28] A. Khraisat, I. Gondal, P. Vamplew, and J. Kamruzzaman, "Survey of intrusion detection systems: techniques, datasets and challenges," *Cybersecurity*, vol. 2, no. 1, p. 20, 2019. [Online]. Available: https://doi.org/10.1186/s42400-019-0038-7

[29] R. Mitchell and I.-R. Chen, "Behavior Rule Based Intrusion Detection for Supporting Secure Medical Cyber Physical Systems," in *2012 21st International Conference on Computer Communications and Networks (ICCCN)*, Jul. 2012, pp. 1–7, iSSN: 1095-2055.

[30] M. Athinaiou, H. Mouratidis, T. Fotis, and M. Pavlidis, "A Conceptual Redesign of a Modelling Language for Cyber Resiliency of Healthcare Systems," in *Computer Security*, ser. Lecture Notes in Computer Science, S. Katsikas, F. Cuppens, N. Cuppens, C. Lambrinoudakis, C. Kalloniatis, J. Mylopoulos, A. Antón, S. Gritzalis, F. Pallas, J. Pohle, A. Sasse, W. Meng, S. Furnell, and

J. Garcia-Alfaro, Eds. Cham: Springer International Publishing, 2020, pp. 140–158.

[31] S. Nandy, M. Adhikari, M. A. Khan, V. G. Menon, and S. Verma, "An Intrusion Detection Mechanism for Secured IoMT framework based on Swarm-Neural Network," *IEEE Journal of Biomedical and Health Informatics*, pp. 1–1, 2021, conference Name: IEEE Journal of Biomedical and Health Informatics.

[32] C. Greer, M. Burns, D. Wollman, and E. Griffor, "Cyber-physical systems and internet of things.(no. special publication (nist sp)-1900-202)," tech. rep, Tech. Rep., 2019.

[33] J. Cohen, "Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit." *Psychological bulletin*, vol. 70, no. 4, p. 213, 1968.

[34] C. Ardito, T. Di Noia, E. Di Sciascio, D. Lofù, A. Pazienza, and F. Vitulano, "User Feedback to Improve the Performance of a Cyber-attack Detection Artificial Intelligence System in the e-Health Domain," in *Human-Computer Interaction – INTERACT 2021*, C. Ardito, R. Lanzilotti, A. Malizia, H. Petrie, A. Piccinno, G. Desolda, and K. Inkpen, Eds. Cham: Springer International Publishing, 2021, vol. 12936, pp. 295–299, series Title: Lecture Notes in Computer Science. [Online]. Available: https://link.springer.com/10.1007/978-3-030-85607-6_25

[35] M. Ahmed, S. Byreddy, A. Nutakki, L. F. Sikos, and P. Haskell-Dowland, "ECU-IoHT: A dataset for analyzing cyberattacks in Internet of Health Things," *Ad Hoc Networks*, vol. 122, p. 102621, Nov. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1570870521001475

[36] F. Akram, D. Liu, P. Zhao, N. Kryvinska, S. Abbas, and M. Rizwan, "Trustworthy Intrusion Detection in E-Healthcare Systems," *Frontiers in Public Health*, vol. 9, p. 788347, Dec. 2021. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8678532/

[37] M. Akshay Kumaar, D. Samiayya, P. M. D. R. Vincent, K. Srinivasan, C.-Y. Chang, and H. Ganesh, "A Hybrid Framework for Intrusion Detection in Healthcare Systems Using Deep Learning," *Frontiers in Public Health*, vol. 9, p. 824898, Jan. 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8790147/

[38] A. Alamleh, O. S. Albahri, A. A. Zaidan, A. S. Albahri, A. H. Alamoodi, B. B. Zaidan, S. Qahtan, H. A. Alsatar, M. S. Al-Samarraay, and A. N. Jasim, "Federated Learning for IoMT Applications: A Standardization and Benchmarking Framework of Intrusion Detection Systems," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 2, pp. 878–887, Feb. 2023, conference Name: IEEE Journal of Biomedical and Health Informatics.

[39] M. A. Almaiah, A. Ali, F. Hajjej, M. F. Pasha, and M. A. Alohali, "A Lightweight Hybrid Deep Learning Privacy Preserving Model for FC-Based Industrial Internet of Medical Things," *Sensors*, vol. 22, no. 6, p. 2112, Mar. 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/6/2112

[40] A. S. Alotaibi, "Biserial Miyaguchi–Preneel Blockchain-Based Ruzicka-Indexed Deep Perceptive Learning for Malware Detection in IoMT," *Sensors*, vol. 21, no. 21, p. 7119, Oct. 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/21/7119

[41] I. Alrashdi, A. Alqazzaz, R. Alharthi, E. Aloufi, M. A. Zohdy, and H. Ming, "FBAD: Fog-based Attack Detection for IoT Healthcare in Smart Cities," in *2019 IEEE 10th Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON)*, Oct. 2019, pp. 0515–0522.

[42] A. Arfaoui, A. Kribeche, S. M. Senouci, and M. Hamdi, "Game-based adaptive anomaly detection in wireless body area networks," *Computer Networks*, vol. 163, p. 106870, Nov. 2019. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1389128619300015

[43] E. Ashraf, N. F. F. Areed, H. Salem, E. H. Abdelhay, and A. Farouk, "FIDChain: Federated Intrusion Detection System for Blockchain-Enabled IoT Healthcare Applications," *Healthcare (Basel, Switzerland)*, vol. 10, no. 6, p. 1110, Jun. 2022.

[44] P. V. Astillo, D. G. Duguma, H. Park, J. Kim, B. Kim, and I. You, "Federated intelligence of anomaly detection agent in IoTMD-enabled Diabetes Management Control System," *Future Generation Computer Systems*, vol. 128, pp. 395–405, Mar. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167739X21004192

[45] J. B. Awotunde, K. M. Abiodun, E. A. Adeniyi, S. O. Folorunso, and R. G. Jimoh, "A Deep Learning-Based Intrusion Detection Technique for a Secured IoMT System," in *Informatics and Intelligent Applications*, ser. Communications in Computer and Information Science, S. Misra, J. Oluranti, R. Damaševičius, and R. Maskeliunas, Eds. Cham: Springer International Publishing, 2022, pp. 50–62.

[46] S.-a. Ayoub, A.-G. Mohammed Ali, and B. Narhimene, "Enhanced Intrusion Detection System for Remote Healthcare," in *Advances in Computing Systems and Applications*, ser. Lecture Notes in Networks and Systems, M. R. Senouci, S. Y. Boulahia, and M. A. Benatia, Eds. Cham: Springer International Publishing, 2022, pp. 323–333.

[47] V. B. Balasubramanyn, G. Thamilarasu, and R. Sridhar, "Security Solution For Data Integrity InWireless BioSensor Networks," in *27th International Conference on Distributed Computing Systems Workshops (ICDCSW'07)*, Jun. 2007, pp. 79–79, iSSN: 1545-0678.

[48] A. Basharat, M. M. B. Mohamad, and A. Khan, "Machine Learning Techniques for Intrusion Detection in Smart Healthcare Systems: A Comparative Analysis," in *2022 4th International Conference on Smart Sensors and Application (ICSSA)*, Jul. 2022, pp. 29–33.

[49] A. Bassene and B. Gueye, "DeepDDoS: A Deep-Learning Model for Detecting Software Defined Healthcare IoT Networks Attacks," in *Ubiquitous Networking*, ser. Lecture Notes in Computer Science, H. Elbiaze, E. Sabir, F. Falcone, M. Sadik, S. Lasaulce, and J. Ben Othman, Eds. Cham: Springer International Publishing, 2021, pp. 201–209.

[50] H. Cai, T. Yun, J. Hester, and K. K. Venkatasubramanian, "Deploying Data-Driven Security Solutions on Resource-Constrained Wearable IoT Systems," in *2017 IEEE 37th International Conference on Distributed Computing Systems Workshops (ICDCSW)*, Jun. 2017, pp. 199–204, iSSN: 2332-5666.

[51] A. S. Carreon-Rascon and J. W. Rozenblit, "Towards Requirements for Self-Healing as a Means of Mitigating Cyber-Intrusions in Medical Devices," in *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Oct. 2022, pp. 1500–1505, iSSN: 2577-1655.

[52] M. Chowdhury, S. Jahan, R. Islam, and J. Gao, "Malware Detection for Healthcare Data Security," in *Security and Privacy in Communication Networks*, ser. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, R. Beyah, B. Chang, Y. Li, and S. Zhu, Eds. Cham: Springer International Publishing, 2018, pp. 407–416.

[53] L. Fernández Maimó, A. Huertas Celdrán, A. L. Perales Gómez, F. J. García Clemente, J. Weimer, and I. Lee, "Intelligent and Dynamic Ransomware Spread Detection and Mitigation in Integrated Clinical Environments," *Sensors*, vol. 19, no. 5, p. 1114, Jan. 2019, number: 5 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/1424-8220/19/5/1114

[54] M. A. Ferrag, O. Friha, L. Maglaras, H. Janicke, and L. Shu, "Federated Deep Learning for Cyber Security in the Internet of Things: Concepts, Applications, and Experimental Analysis," *IEEE Access*, vol. 9, pp. 138 509–138 542, 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9562531/

[55] M. Fouda, R. Ksantini, and W. Elmedany, "A Novel Intrusion Detection System for Internet of Healthcare Things Based on Deep Subclasses Dispersion Information," *IEEE Internet of Things Journal*, pp. 1–1, 2022, conference Name: IEEE Internet of Things Journal.

[56] S. Gao and G. Thamilarasu, "Machine-Learning Classifiers for Security in Connected Medical Devices," in *2017 26th International Conference on Computer Communication and Networks (ICCCN)*. Vancouver, BC, Canada: IEEE, Jul. 2017, pp. 1–5. [Online]. Available: http://ieeexplore.ieee.org/document/8038507/

[57] A. Ghourabi, "A Security Model Based on LightGBM and Transformer to Protect Healthcare Systems From Cyberattacks," *IEEE Access*, vol. 10, pp. 48 890–48 903, 2022, conference Name: IEEE Access.

[58] D. Gupta, M. Gupta, S. Bhatt, and A. S. Tosun, "Detecting Anomalous User Behavior in Remote Patient Monitoring," in *2021 IEEE 22nd International Conference on Information Reuse and Integration for Data Science (IRI)*, Aug. 2021, pp. 33–40.

[59] D. Gupta, O. Kayode, S. Bhatt, M. Gupta, and A. S. Tosun, "Hierarchical Federated Learning based Anomaly Detection using Digital Twins for Smart Healthcare," *arXiv:2111.12241 [cs]*, Nov. 2021, arXiv: 2111.12241. [Online]. Available: http://arxiv.org/abs/2111.12241

[60] K. Gupta, D. K. Sharma, K. Datta Gupta, and A. Kumar, "A tree classifier based network intrusion detection model for Internet of Medical Things," *Computers and Electrical Engineering*, vol. 102, Sep. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045790622004049

[61] A. A. Hady, A. Ghubaish, T. Salman, D. Unal, and R. Jain, "Intrusion Detection System for Healthcare Systems Using Medical and Network Data: A Comparison Study," *IEEE Access*, vol. 8, pp. 106 576–106 584, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9109651/

[62] M. Hajder, J. Kolbusz, P. Hajder, M. Nycz, and M. Liput, "Data Security Platform Model in Networked Medical IT Systems based on Statistical Classifiers and ANN," *Procedia Computer Science*, vol. 176, pp. 3682–3691, 2020. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1877050920319104

[63] S. S. Hameed, W. H. Hassan, and L. A. Latiff, "An Efficient Fog-Based Attack Detection Using Ensemble of MOA-WMA for Internet of Medical Things," in *Innovative Systems for Intelligent Health Informatics*, ser. Lecture Notes on Data Engineering and Communications Technologies, F. Saeed, F. Mohammed, and A. Al-Nahari, Eds. Cham: Springer International Publishing, 2021, pp. 774–785.

[64] S. S. Hameed, A. Selamat, L. Abdul Latiff, S. A. Razak, O. Krejcar, H. Fujita, M. N. Ahmad Sharif, and S. Omatu, "A Hybrid Lightweight System for Early Attack Detection in the IoMT Fog," *Sensors*, vol. 21, no. 24, p. 8289, Dec. 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/24/8289

[65] S. S. Hameed, A. Selamat, L. A. Latiff, S. A. Razak, and O. Krejcar, "WHTE: Weighted Hoeffding Tree Ensemble for Network Attack Detection at Fog-IoMT," in *Advances and Trends in Artificial Intelligence. Theory and Practices in Artificial Intelligence*, ser. Lecture Notes in Computer Science, H. Fujita, P. Fournier-Viger, M. Ali, and Y. Wang, Eds. Cham: Springer International Publishing, 2022, pp. 485–496.

[66] N. I. Haque, A. A. Khalil, M. A. Rahman, M. H. Amini, and S. I. Ahamed, "BIOCAD: Bio-Inspired Optimization for Classification and Anomaly Detection in Digital Healthcare Systems," in *2021 IEEE International Conference on Digital Health (ICDH)*, Sep. 2021, pp. 48–58.

[67] N. I. Haque, M. A. Rahman, and S. I. Ahamed, "DeepCAD: A Stand-alone Deep Neural Network-based Framework for Classification and Anomaly Detection in Smart Healthcare Systems," in *2022 IEEE International Conference on Digital Health (ICDH)*, Jul. 2022, pp. 218–227.

[68] D. He, Q. Qiao, Y. Gao, J. Zheng, S. Chan, J. Li, and N. Guizani, "Intrusion Detection Based on Stacked Autoencoder for Connected Healthcare Systems," *IEEE Network*, vol. 33, no. 6, pp. 64–69, Nov. 2019, conference Name: IEEE Network.

[69] X. Hei, X. Du, S. Lin, I. Lee, and O. Sokolsky, "Patient Infusion Pattern based Access Control Schemes for Wireless Insulin Pump System," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 11, pp. 3108–3121, Nov. 2015.

[70] F. Hussain, S. G. Abbas, G. A. Shah, I. M. Pires, U. U. Fayyaz, F. Shahzad, N. M. Garcia, and E. Zdravevski, "A Framework for Malicious Traffic Detection in IoT Healthcare Environment," *Sensors*, vol. 21, no. 9, p. 3025, Apr. 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/9/3025

[71] O. Igbe, I. Darwish, and T. Saadawi, "Distributed Network Intrusion Detection Systems: An Artificial Immune System Approach," in *2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, Jun. 2016, pp. 101–106.

[72] M. J. Iqbal, S. Aurangzeb, M. Aleem, G. Srivastava, and J. C.-W. Lin, "RThreatDroid: A Ransomware Detection Approach to Secure IoT Based Healthcare Systems," *IEEE Transactions on Network Science and Engineering*, pp. 1–10, 2022, conference Name: IEEE Transactions on Network Science and Engineering.

[73] A. Karthick Kumar, K. Vadivukkarasi, R. Dayana, and P. Malarvezhi, "Botnet Attacks Detection Using Embedded Feature Selection Methods for Secure IOMT Environment," in *Pervasive Computing and Social Networking*, ser. Lecture Notes in Networks and Systems, G. Ranganathan, R. Bestak, and X. Fernando, Eds. Singapore: Springer Nature, 2023, pp. 585–599.

[74] I. A. Khan, N. Moustafa, I. Razzak, M. Tanveer, D. Pi, Y. Pan, and B. S. Ali, "XSRU-IoMT: Explainable simple recurrent units for threat detection in Internet of Medical Things networks," *Future Generation Computer Systems*, vol. 127, pp. 181–193, Feb. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167739X21003563

[75] F. Khan, M. A. Jan, R. Alturki, M. D. Alshehri, S. T. Shah, and A. u. Rehman, "A Secure Ensemble Learning-Based Fog-Cloud Approach for Cyberattack Detection in IoMT," *IEEE Transactions on Industrial Informatics*, pp. 1–9, 2023, conference Name: IEEE Transactions on Industrial Informatics.

[76] I. Firat Kilincer, F. Ertam, A. Sengur, R.-S. Tan, and U. Rajendra Acharya, "Automated detection of cybersecurity attacks in healthcare systems with recursive feature elimination and multilayer perceptron optimization," *Biocybernetics and Biomedical Engineering*, vol. 43, no. 1, Jan. 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0208521622001012

[77] M. Kintzlinger, A. Cohen, N. Nissim, M. Rav-Acha, V. Khalameizer, Y. Elovici, Y. Shahar, and A. Katz, "CardiWall: A Trusted Firewall for the Detection of Malicious Clinical Programming of Cardiac Implantable Electronic Devices," *IEEE Access*, vol. 8, pp. 48 123–48 140, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9025056/

[78] P. Kumar, G. P. Gupta, and R. Tripathi, "An ensemble learning and fog-cloud architecture-driven cyber-attack detection framework for IoMT networks," *Computer Communications*, vol. 166, pp. 110–124, Jan. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0140366420320090

[79] R. Kumar, P. Kumar, R. Tripathi, G. P. Gupta, A. K. M. N. Islam, and M. Shorfuzzaman, "Permissioned Blockchain and Deep Learning for Secure and Efficient Data Sharing in Industrial Healthcare Systems," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 11, pp. 8065–8073, Nov. 2022, conference Name: IEEE Transactions on Industrial Informatics.

[80] P. Kumar, R. Kumar, S. Garg, K. Kaur, Y. Zhang, and M. Guizani, "A Secure Data Dissemination Scheme for IoT-Based e-Health Systems using AI and Blockchain," in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, Dec. 2022, pp. 1397–1403.

[81] W. Li, W. Meng, C. Su, and L. F. Kwok, "Towards False Alarm Reduction Using Fuzzy If-Then Rules for Medical Cyber Physical Systems," *IEEE Access*, vol. 6, pp. 6530–6539, 2018, conference Name: IEEE Access.

[82] T. Mahler, E. Shalom, Y. Elovici, and Y. Shahar, "A dual-layer context-based architecture for the detection of anomalous instructions sent to medical devices," *Artificial Intelligence in Medicine*, vol. 123, p. 102229, Jan. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0933365721002220

[83] S. Manimurugan, S. Al-Mutairi, M. M. Aborokbah, N. Chilamkurti, S. Ganesan, and R. Patan, "Effective Attack Detection in Internet of Medical Things Smart Environment Using a Deep Belief Neural Network," *IEEE Access*, vol. 8, pp. 77 396–77 404, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9057709/

[84] A. Mishra and P. Bagade, "Digital Forensics for Medical Internet of Things," in *2022 IEEE Globecom Workshops (GC Wkshps)*, Dec. 2022, pp. 1074–1079.

[85] R. Anwar Mohammed and K. M. Ali Alheeti, "Intrusion detection system for Healthcare based on Convolutional Neural Networks," in *2022 Iraqi International Conference on Communication and Information Technologies (IICCIT)*, Sep. 2022, pp. 216–221.

[86] N. Mowla, I. Doh, and K. Chae, "Evolving neural network intrusion detection system for MCPS," in *2018 20th International Conference on Advanced Communication Technology (ICACT)*, Feb. 2018, pp. 1040–1045.

[87] T. Muhammed, R. Mehmood, A. Albeshri, and I. Katib, "UbeHealth: A Personalized Ubiquitous Cloud and Edge-Enabled Networked Healthcare System for Smart Cities," *IEEE Access*, vol. 6, pp. 32 258–32 285, 2018, conference Name: IEEE Access.

[88] J. Nayak, S. K. Meher, A. Souri, B. Naik, and S. Vimal, "Extreme learning machine and bayesian optimization-driven intelligent framework for IoMT cyber-attack detection," *The Journal of Supercomputing*, vol. 78, no. 13, pp. 14 866–14 891, 2022.

[89] A. I. Newaz, A. K. Sikder, M. A. Rahman, and A. S. Uluagac, "HealthGuard: A Machine Learning-Based Security Framework for Smart Healthcare Systems," in *2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, Oct. 2019, pp. 389–396.

[90] A. I. Newaz, A. K. Sikder, L. Babun, and A. S. Uluagac, "HEKA: A Novel Intrusion Detection System for Attacks to Personal Medical Devices," in *2020 IEEE Conference on Communications and Network Security (CNS)*, Jun. 2020, pp. 1–9.

[91] Y. Otoum, Y. Wan, and A. Nayak, "Federated Transfer Learning-Based IDS for the Internet of Medical Things (IoMT)," in *2021 IEEE Globecom Workshops (GC Wkshps)*, Dec. 2021, pp. 1–6.

[92] S. Otoum, N. Guizani, and H. Mouftah, "Federated Reinforcement Learning-Supported IDS for IoT-steered Healthcare Systems," in *ICC 2021 - IEEE International Conference on Communications*, Jun. 2021, pp. 1–6, iSSN: 1938-1883.

[93] Y. Otoum, V. Chamola, and A. Nayak, "Federated and Transfer Learning-Empowered Intrusion Detection for IoT Applications," *IEEE Internet of Things Magazine*, vol. 5, no. 3, pp. 50–54, Sep. 2022, conference Name: IEEE Internet of Things Magazine.

[94] D. Panagoda, C. Malinda, C. Wijetunga, L. Rupasinghe, B. Bandara, and C. Liyanapathirana, "Application of Federated Learning in Health Care Sector for Malware Detection and Mitigation Using Software Defined Networking Approach," in *2022 2nd Asian Conference on Innovation in Technology (ASIANCON)*, Aug. 2022, pp. 1–6.

[95] S. P. R.m., P. K. R. Maddikunta, P. M., S. Koppu, T. R. Gadekallu, C. L. Chowdhary, and M. Alazab, "An effective feature engineering for DNN using hybrid PCA-GWO for intrusion detection in IoMT architecture," *Computer Communications*, vol. 160, pp. 139–149, Jul. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S014036642030298X

[96] P. Radoglou-Grammatikis, P. Sarigiannidis, G. Efstathopoulos, T. Lagkas, G. Fragulis, and A. Sarigiannidis, "A Self-Learning Approach for Detecting Intrusions in Healthcare Systems," in *ICC 2021 - IEEE International Conference on Communications*, Jun. 2021, pp. 1–6, iSSN: 1938-1883.

[97] P. Radoglou-Grammatikis, K. Rompolos, P. Sarigiannidis, V. Argyriou, T. Lagkas, A. Sarigiannidis, S. Goudos, and S. Wan, "Modeling, Detecting, and Mitigating Threats Against Industrial Healthcare Systems: A Combined Software Defined Networking and Reinforcement Learning Approach," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 3, pp. 2041–2052, Mar. 2022, conference Name: IEEE Transactions on Industrial Informatics.

[98] S. Rahmadika, P. V. Astillo, G. Choudhary, D. G. Duguma, V. Sharma, and I. You, "Blockchain-based Privacy Preservation Scheme for Misbehavior Detection in Lightweight IoMT Devices," *IEEE journal of biomedical and health informatics*, vol. PP, Jun. 2022.

[99] N. Ram and D. Kumar, "Effective Cyber Attack Detection in an IoMT-Smart System using Deep Convolutional Neural Networks and Machine Learning Algorithms," in *2022 Second International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering (ICATIECE)*, Dec. 2022, pp. 1–6.

[100] A. Rao, N. Carreon, R. Lysecky, and J. Rozenblit, "Probabilistic Threat Detection for Risk Management in Cyber-physical Medical Systems," *IEEE Software*, vol. 35, no. 1, pp. 38–43, Jan. 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8239935/

[101] A. Rao, N. Carreón, R. Lysecky, J. Rozenblit, and J. Sametinger, "Resilient Security of Medical Cyber-Physical Systems," in *Database and Expert Systems Applications*, ser. Communications in Computer and Information Science, G. Anderst-Kotsis, A. M. Tjoa, I. Khalil, M. Elloumi, A. Mashkoor, J. Sametinger, X. Larrucea, A. Fensel, J. Martinez-Gil, B. Moser, C. Seifert, B. Stein, and M. Granitzer, Eds. Cham: Springer International Publishing, 2019, pp. 95–100.

[102] V. Ravi, T. D. Pham, and M. Alazab, "Attention-Based Multidimensional Deep Learning Approach for Cross-Architecture IoMT Malware Detection and Classification in Healthcare Cyber-Physical Systems," *IEEE Transactions on Computational Social Systems*, pp. 1–10, 2022, conference Name: IEEE Transactions on Computational Social Systems.

[103] A. Rehman, S. Abbas, M. A. Khan, T. M. Ghazal, K. M. Adnan, and A. Mosavi, "A secure healthcare 5.0 system based on blockchain technology entangled with federated learning technique," *Computers in Biology and Medicine*, vol. 150, p. 106019, Sep. 2022.

[104] T. Saba, "Intrusion Detection in Smart City Hospitals using Ensemble Classifiers," in *2020 13th International Conference on Developments in eSystems Engineering (DeSE)*, Dec. 2020, pp. 418–422, iSSN: 2161-1351.

[105] Y. K. Saheed and M. O. Arowolo, "Efficient Cyber Attack Detection on the Internet of Medical Things-Smart Environment Based on Deep Recurrent Neural Network and Machine Learning Algorithms," *IEEE Access*, vol. 9, pp. 161 546–161 554, 2021, conference Name: IEEE Access.

[106] A. M. Said, A. Yahyaoui, F. Yaakoubi, and T. Abdellatif, "Machine learning based rank attack detection for smart hospital infrastructure," in *International Conference on Smart Homes and Health Telematics*. Springer, 2020, pp. 28–40.

[107] O. Salem and A. Mehaoua, "A Secure Framework for Remote Healthcare Monitoring using the Internet of Medical Things," in *ICC 2022 - IEEE International Conference on Communications*, May 2022, pp. 1233–1238, iSSN: 1938-1883.

[108] W. Schneble and G. Thamilarasu, "Optimal Feature Selection for Intrusion Detection in Medical Cyber-Physical Systems," in *2019 11th International Conference on Advanced Computing (ICoAC)*, Dec. 2019, pp. 238–243.

[109] N. Sehatbakhsh, M. Alam, A. Nazari, A. Zajic, and M. Prvulovic, "Syndrome: Spectral analysis for anomaly detection on medical IoT and embedded devices," in *2018 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*, Apr. 2018, pp. 1–8.

[110] R. A. Sharma, I. Sabane, M. Apostolaki, A. Rowe, and V. Sekar, "Lumen: a framework for developing and evaluating ML-based IoT network anomaly detection," in *Proceedings of the 18th International Conference on emerging Networking EXperiments and Technologies*, ser. CoNEXT '22. New York, NY, USA: Association for Computing Machinery, Nov. 2022, pp. 59–71. [Online]. Available: https://doi.org/10.1145/3555050.3569129

[111] P. Singh, G. S. Gaba, A. Kaur, M. Hedabou, and A. Gurtov, "Dew-Cloud-Based Hierarchical Federated Learning for Intrusion Detection in IoMT," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 2, pp. 722–731, Feb. 2023, conference Name: IEEE Journal of Biomedical and Health Informatics.

[112] I. Siniosoglou, P. Sarigiannidis, V. Argyriou, T. Lagkas, S. K. Goudos, and M. Poveda, "Federated Intrusion Detection In NG-IoT Healthcare Systems: An Adversarial Approach," in *ICC 2021 - IEEE International Conference on Communications*, Jun. 2021, pp. 1–6, iSSN: 1938-1883.

[113] F. Spegni, A. Sabatelli, A. Merlo, L. Pepa, L. Spalazzi, and L. Verderame, "A Precision Cybersecurity Workflow for Cyber-physical Systems: The IoT Healthcare Use Case," in *Computer Security. ESORICS 2022 International Workshops*, ser. Lecture Notes in Computer Science, S. Katsikas, F. Cuppens, C. Kalloniatis, J. Mylopoulos, F. Pallas, J. Pohle, M. A. Sasse, H. Abie, S. Ranise, L. Verderame, E. Cambiaso, J. Maestre Vidal, M. A. Sotelo Monge, M. Albanese, B. Katt, S. Pirbhulal, and A. Shukla, Eds. Cham: Springer International Publishing, 2023, pp. 409–426.

[114] A. Tabassum, A. Erbad, A. Mohamed, and M. Guizani, "Privacy-Preserving Distributed IDS Using Incremental Learning for IoT Health Systems," *IEEE Access*, vol. 9, pp. 14 271–14 283, 2021, conference Name: IEEE Access.

[115] B. Tahir, A. Jolfaei, and M. Tariq, "A Novel Experience-Driven and Federated Intelligent Threat-Defense Framework in IoMT," *IEEE Journal of Biomedical and Health Informatics*, pp. 1–8, 2023, conference Name: IEEE Journal of Biomedical and Health Informatics.

[116] G. Thamilarasu, A. Odesile, and A. Hoang, "An Intrusion Detection System for Internet of Medical Things," *IEEE Access*, vol. 8, pp. 181 560–181 576, 2020, conference Name: IEEE Access.

[117] C. Thapa, K. K. Karmakar, A. H. Celdran, S. Camtepe, V. Varadharajan, and S. Nepal, "FedDICE: A ransomware spread detection in a distributed integrated clinical environment using federated learning and SDN based mitigation," *arXiv:2106.05434 [cs]*, vol. 402, pp. 3–24, 2021, arXiv: 2106.05434. [Online]. Available: http://arxiv.org/abs/2106.05434

[118] A. A. Toor, M. Usman, F. Younas, A. C. M. Fong, S. A. Khan, and S. Fong, "Mining Massive E-Health Data Streams for IoMT Enabled Healthcare Systems," *Sensors*, vol. 20, no. 7, p. 2131, Apr. 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/7/2131

[119] S. Masengo Wa Umba, A. M. Abu-Mahfouz, and D. Ramotsoela, "Artificial Intelligence-Driven Intrusion Detection in Software-Defined Wireless Sensor Networks: Towards Secure IoT-Enabled Healthcare Systems," *International Journal of Environmental Research and Public Health*, vol. 19, no. 9, p. 5367, Apr. 2022.

[120] S. A. Wagan, J. Koo, I. F. Siddiqui, N. M. F. Qureshi, M. Attique, and D. R. Shin, "A Fuzzy-Based Duo-Secure Multi-Modal Framework for IoMT Anomaly Detection," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 1, pp. 131–144, Jan. 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1319157822004050

[121] F. Wahab, Y. Zhao, D. Javeed, M. H. Al-Adhaileh, S. A. Almaaytah, W. Khan, M. S. Saeed, and R. Kumar Shah, "An AI-Driven Hybrid Framework for Intrusion Detection in IoT-Enabled E-Health," *Computational Intelligence and Neuroscience*, vol. 2022, p. 6096289, 2022.

[122] J. Wang, H. Jin, J. Chen, J. Tan, and K. Zhong, "Anomaly detection in Internet of medical Things with Blockchain from the perspective of deep neural network," *Information Sciences*, vol. 617, Dec. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0020025522011835

[123] R. Yan, V. C. Shah, T. Xu, and M. Potkonjak, "Security Defenses for Vulnerable Medical Sensor Network," in *2014 IEEE International Conference on Healthcare Informatics*, Sep. 2014, pp. 300–309.

[124] B. Zaabar, O. Cheikhrouhou, and M. Abid, "Intrusion Detection System for IoMT through Blockchain-based Federated Learning," in *2022 15th International Conference on Security of Information and Networks (SIN)*, Nov. 2022, pp. 01–08.

[125] G. Zachos, G. Mantas, I. Essop, K. Porfyrakis, J. C. Ribeiro, and J. Rodriguez, "Prototyping an Anomaly-Based Intrusion Detection System for Internet of Medical Things Networks," in *2022 IEEE 27th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, Nov. 2022, pp. 179–183, iSSN: 2378-4873.

[126] M. Zubair, A. Ghubaish, D. Unal, A. Al-Ali, T. Reimann, G. Alinier, M. Hammoudeh, and J. Qadir, "Secure Bluetooth Communication in Smart Healthcare Systems: A Novel Community Dataset and Intrusion Detection System," *Sensors (Basel, Switzerland)*, vol. 22, no. 21, p. 8280, Oct. 2022.

[127] R. Abdulhammed, M. Faezipour, and K. Elleithy, "Malicious behavior monitoring of embedded medical devices," in *2017 IEEE Long Island Systems, Applications and Technology Conference (LISAT)*, May 2017, pp. 1–6.

[128] G. Choudhary, P. V. Astillo, I. You, K. Yim, I.-R. Chen, and J.-H. Cho, "Lightweight Misbehavior Detection Management of Embedded IoT Devices in Medical Cyber Physical Systems," *IEEE Transactions on Network and Service Management*, vol. 17, no. 4, pp. 2496–2510, Dec. 2020, conference Name: IEEE Transactions on Network and Service Management.

[129] L. Fang, Y. Li, Z. Liu, C. Yin, M. Li, and Z. J. Cao, "A Practical Model Based on Anomaly Detection for Protecting Medical IoT Control Services Against External Attacks," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 6, pp. 4260–4269, Jun. 2021, conference Name: IEEE Transactions on Industrial Informatics.

[130] R. Mitchell and I.-R. Chen, "Behavior Rule Specification-Based Intrusion Detection for Safety Critical Medical Cyber Physical Systems," *IEEE Transactions on Dependable and Secure Computing*, vol. 12, no. 1, pp. 16–30, Jan. 2015, conference Name: IEEE Transactions on Dependable and Secure Computing.

[131] Z. Li, L. Cheng, and Y. Zhang, "Tracking Sensitive Information and Operations in Integrated Clinical Environment," in *2019 18th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/13th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*, Aug. 2019, pp. 192–199, iSSN: 2324-9013.

[132] M. Raiyat Aliabadi, M. Seltzer, M. Vahidi Asl, and R. Ghavamizadeh, "ARTINALI#: An Efficient Intrusion Detection Technique for Resource-Constrained Cyber-Physical Systems," *International Journal of Critical Infrastructure Protection*, vol. 33, p. 100430, Jun. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1874548221000226

[133] I. You, K. Yim, V. Sharma, G. Choudhary, I.-R. Chen, and J.-H. Cho, "Misbehavior detection of embedded IoT devices in medical cyber physical systems," in *Proceedings of the 2018 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies*. Washington DC: ACM, Sep. 2018, pp. 88–93. [Online]. Available: https://dl.acm.org/doi/10.1145/3278576.3278601

[134] M. Zhang, A. Raghunathan, and N. K. Jha, "MedMon: Securing Medical Devices Through Wireless Monitoring and Anomaly Detection," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 7, no. 6, Dec. 2013.

[135] M. Boujrad, S. Lazaar, and M. Hassine, "Performance Assessment of Open Source IDS for improving IoT Architecture Security implemented on WBANs," in *Proceedings of the 3rd International Conference on Networking, Information Systems & Security*. Marrakech Morocco: ACM, Mar. 2020, pp. 1–4. [Online]. Available: https://dl.acm.org/doi/10.1145/3386723.3387892

[136] C. Mpungu, C. George, and G. Mapp, "Developing a Novel Digital Forensics Readiness Framework for Wireless Medical Networks Using Specialised Logging," in *Cybersecurity in the Age of Smart Societies*, ser. Advanced Sciences and Technologies for Security Applications, H. Jahankhani, Ed. Cham: Springer International Publishing, 2023, pp. 203–226.

[137] H. Zhang, S. Kang, and Y. Li, "Visual Construction Algorithm of Attack Path Based on Medical Sensor Networks," in *2018 IEEE International Conference of Safety Produce Informatization (IICSPI)*, Dec. 2018, pp. 775–779.

[138] M. Begli, F. Derakhshan, and H. Karimipour, "A Layered Intrusion Detection System for Critical Infrastructure Using Machine Learning," in *2019 IEEE 7th International Conference on Smart Energy Grid Engineering (SEGE)*, Aug. 2019, pp. 120–124, iSSN: 2575-2693.

[139] M. Chen, Y. Qian, J. Chen, K. Hwang, S. Mao, and L. Hu, "Privacy Protection and Intrusion Avoidance for Cloudlet-Based Medical Data Sharing," *IEEE Transactions on Cloud Computing*, vol. 8, no. 4, pp. 1274–1283, Oct. 2020, conference Name: IEEE Transactions on Cloud Computing.

[140] G. Dupont, D. R. dos Santos, E. Costante, J. den Hartog, and S. Etalle, "A Matter of Life and Death: Analyzing the Security of Healthcare Networks," in *ICT Systems Security and Privacy Protection*, ser. IFIP Advances in Information and Communication Technology, M. Hölbl, K. Rannenberg, and T. Welzer, Eds. Cham: Springer International Publishing, 2020, pp. 355–369.

[141] W. Meng, K.-K. R. Choo, S. Furnell, A. V. Vasilakos, and C. W. Probst, "Towards Bayesian-Based Trust Management for Insider Attacks in Healthcare Software-Defined Networks," *IEEE Transactions on Network and Service Management*, vol. 15, no. 2, pp. 761–773, Jun. 2018, conference Name: IEEE Transactions on Network and Service Management.

[142] N. Kolokotronis, M. Dareioti, S. Shiaeles, and E. Bellini, "An Intelligent Platform for Threat Assessment and Cyber-Attack Mitigation in IoMT Ecosystems," in *2022 IEEE Globecom Workshops (GC Wkshps)*, Dec. 2022, pp. 541–546.

[143] E. Lakka, G. Hatzivasilis, S. Karagiannis, A. Alexopoulos, M. Athanatos, S. Ioannidis, M. Chatzimpyrros, G. Kalogiannis, and G. Spanoudakis, "Incident Handling for Healthcare Organizations and Supply-Chains," in *2022 IEEE Symposium on Computers and Communications (ISCC)*, Jun. 2022, pp. 1–7, iSSN: 2642-7389.

[144] U. Tariq, I. Ullah, M. Yousuf Uddin, and S. J. Kwon, "An Effective Self-Configurable Ransomware Prevention Technique for IoMT," *Sensors (Basel, Switzerland)*, vol. 22, no. 21, p. 8516, Nov. 2022.

[145] H. Nguyen, B. Acharya, R. Ivanov, A. Haeberlen, L. T. X. Phan, O. Sokolsky, J. Walker, J. Weimer, W. Hanson, and I. Lee, "Cloud-Based Secure Logger for Medical Devices," in *2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, Jun. 2016, pp. 89–94.

[146] S. D. Hettich, S Bay, "Kdd cup 1999 data," *The UCI KDD Archive*, 1999.

[147] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*. Ottawa, ON, Canada: IEEE, Jul. 2009, pp. 1–6. [Online]. Available: http://ieeexplore.ieee.org/document/5356528/

[148] I. Sharafaldin, A. Gharib, A. H. Lashkari, and A. A. Ghorbani, "Towards a reliable intrusion detection benchmark dataset," *Software Networking*, vol. 2018, no. 1, pp. 177–200, 2018.

[149] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization." *ICISSp*, vol. 1, pp. 108–116, 2018.

[150] A. Alsaedi, N. Moustafa, Z. Tari, A. Mahmood, and A. Anwar, "Ton_iot telemetry dataset: A new generation dataset of iot and iiot for data-driven intrusion detection systems," *IEEE Access*, vol. 8, pp. 165 130–165 150, 2020.

[151] N. Koroniotis, N. Moustafa, E. Sitnikova, and B. Turnbull, "Towards the development of realistic botnet dataset in the internet of things for network

forensic analytics: Bot-iot dataset," *Future Generation Computer Systems*, vol. 100, pp. 779–796, 2019.

[152] A. E. Johnson, T. J. Pollard, L. Shen, L.-w. H. Lehman, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. Anthony Celi, and R. G. Mark, "Mimic-iii, a freely accessible critical care database," *Scientific data*, vol. 3, no. 1, pp. 1–9, 2016.

[153] A. Boddy, W. Hurst, M. Mackay, and A. E. Rhalibi, "A Study into Detecting Anomalous Behaviours within HealthCare Infrastructures," in *2016 9th International Conference on Developments in eSystems Engineering (DeSE)*, Aug. 2016, pp. 111–117, iSSN: 2161-1343.

[154] D. Arney, J. Plourde, and J. M. Goldman, "Openice medical device interoperability platform overview and requirement analysis," *Biomedical Engineering/Biomedizinische Technik*, vol. 63, no. 1, pp. 39–47, 2018.

[155] H. Nguyen-An, T. Silverston, T. Yamazaki, and T. Miyoshi, "Generating IoT traffic: A Case Study on Anomaly Detection," in *2020 IEEE International Symposium on Local and Metropolitan Area Networks (LANMAN*, Jul. 2020, pp. 1–6, iSSN: 1944-0375.

[156] S. S. Clark and K. Fu, "Recent Results in Computer Security for Medical Devices," in *Wireless Mobile Communication and Healthcare*, ser. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, K. S. Nikita, J. C. Lin, D. I. Fotiadis, and M.-T. Arredondo Waldmeyer, Eds. Berlin, Heidelberg: Springer, 2012, pp. 111–118.

Chapter 5

Don't Stop Believin': A Unified Evaluation Approach for LLM Honeypots

This chapter gives an overview of the contributions and the impact of Weber et al. (2024b)[1]:

Simon Weber, Marc Feger, and Michael Pilgermann

"Don't Stop Believin': A Unified Evaluation Approach for LLM Honeypots"

In: *IEEE Access*, Volume 12, pages 144579 - 144587, Oktober 2024

Acceptance Rate: ~27%

5.1 Summary

This paper examines the integration of LLMs, specifically GPT-3.5, as a backend for honeypots to simulate realistic command-response interactions with attackers. Traditional high-interaction honeypots, designed to replicate real operating systems, are accompanied by challenges such as high development costs, vulnerability to breakout risks, and limited versatility in mimicking specialized systems. By leveraging LLMs, these honeypots can enhance interaction realism without the complexities and risks of an actual OS environment.

The study analyzed over 1,400 request-response pairs from three datasets with different complexities, each designed to evaluate the LLM's effectiveness in generating responses that mimic authentic Secure Shell (SSH) servers. A key finding revealed that while GPT-3.5 exhibits limitations in maintaining session context, incorporating contextual information can improve response convincingness. The paper also discusses the reliability of distinguishing between convincing and non-convincing responses through a paraphrase-mining approach, using cosine distance, which achieved a macro F1 score of 77.85%. This metric-based approach aids in minimizing annotation efforts and supports a unified evaluation framework for LLM-based honeypots, enabling comparative studies.

The authors highlight that certain commands, especially those with complex structures or multiple special characters, pose challenges for the LLM, resulting in less convincing responses. The variability in convincingness across datasets points out the importance of dataset composition and session history for LLM performance. The authors conclude that while current LLMs are not yet robust enough to function as flawless high-interaction honeypots, their performance could be enhanced with fine-tuning. The study opens avenues for future research in developing more advanced, context-aware LLM-based honeypots capable of sustaining long-term engagement with attackers, particularly in specific fields, such as hospitals, where fidelity is essential.

5.2 Personal Contribution

Simon Weber had the initial idea for an evaluation of LLMs as honeypot backends, selected and prepared the datasets, and supplemented them with responses from the SSH server and the LLM. Simon Weber and Marc Feger collaboratively developed and tested the annotation framework. Marc Feger led the evaluation of the annotations. Simon Weber conducted the dataset comparison and analyzed command complexity, while Marc Feger assessed transition probabilities and developed the distance-based method for evaluating response convincingness. Simon Weber drafted the initial manuscript, which he and Marc Feger subsequently revised together. Michael Pilgermann contributed by discussing ideas with the authors and providing feedback on the drafts.

5.3 Importance and Impact on this Thesis

Honeypots can be an effective tool for detecting attacks on devices, for which neither host- nor network-based monitoring alone is sufficient. As discussed in section 3.2, up to this date only few robust, publicly available industry-specific honeypots exist. Challenges persist in developing specialized honeypots for critical healthcare equipment, such as MCPS, using traditional methods. LLMs present a promising alternative, given their emerging research focus and unique ability to engage in conversational interactions that convincingly emulate real-world responses. Since network protocols often resemble technical dialogues, LLMs could be especially suited for deceiving attackers into believing they are interacting with authentic devices. Leveraging the broad, pre-trained knowledge base of these models allows for the rapid and cost-effective emulation of various device types, enabling extended attack detection across many device classes and facilitating research into attack trends on these systems. However, this research reveals that a current pre-trained model falls short in generating convincingly realistic responses. The unified evaluation approach provides a structured methodology to enable comparisons of new approaches with pre-trained or fine-tuned models, paving the way towards developing the first fully functional LLM-based honeypot that enhances hospital security and attack detection.

Don't Stop Believin': A Unified Evaluation Approach for LLM Honeypots

SIMON B. WEBER
Heinrich-Heine-University
Düsseldorf, Germany
Simon.Weber@hhu.de

MARC FEGER
Heinrich-Heine-University
Düsseldorf, Germany
Marc.Feger@hhu.de

MICHAEL PILGERMANN
University of Applied Sciences
Brandenburg a. d. Havel, Germany
Michael.Pilgermann@th-brandenburg.de

Abstract

The research area of honeypots is gaining new momentum, driven by advancements in large language models (LLMs). The chat-based applications of generative pretrained transformer (GPT) models seem ideal for the use as honeypot backends, especially in request-response protocols like Secure Shell (SSH). By leveraging LLMs, many challenges associated with traditional honeypots – such as high development costs, ease of exposure, and breakout risks – appear to be solved. While early studies have primarily focused on the potential of these models, our research investigates the current limitations of GPT-3.5 by analyzing three datasets of varying complexity. We conducted an expert annotation of over 1,400 request-response pairs, encompassing 230 different base commands. Our findings reveal that while GPT-3.5 struggles to maintain context, incorporating session context into response generation improves the quality of SSH responses. Additionally, we explored whether distinguishing between convincing and non-convincing responses is a metrics issue. We propose a paraphrase-mining approach to address this challenge, which achieved a macro F1 score of 77.85% using cosine distance in our evaluation. This method has the potential to reduce annotation efforts, converge LLM-based honeypot performance evaluation, and facilitate comparisons between new and previous approaches in future research.

Index Terms

IT Security, Honeypot, Large Language Model, GPT, Cosine Distance, Evaluation

1 Introduction

Honeypots are widely used in research to analyze attacker behavior and in industry contexts to detect and prevent attacks. For several protocols, medium or high-interaction honeypots are prevalent and usable. However, developing high-interaction honeypots is a tedious task, especially if the honeypot should mimic a specific device or service. In addition, attackers often specialize in honeypot detection and develop tests to identify whether a listening service is a honeypot [1]. As a result, using publicly available or well-known honeypots can alienate sophisticated attackers.

Large Language Models (LLMs) have recently gained recognition as valuable tools across various domains, with IT security being a newly emerging area of application [2–4]. Evaluations have shown that ChatGPT (GPT-3.5 and GPT-4) are capable of contextualizing conversations, producing source code, and generating other machine output [5–7]. Despite the diversity of data sources, GPT-3's training dataset includes common web data, among others, which encompass technical manuals, forums, and coding resources [5, 8], suggesting their applicability in honeypot systems.

Since server attacks often resemble technical conversations, with attackers sending commands and expecting specific responses, LLMs might be well-suited to mimic these exchanges and enhance honeypot interaction. With their ability to understand context and generate a wide variety of applicable responses, they could be remarkably effective when an attacker may input unexpected or uncommon commands.

In particular, protocols with a request-response format, such as SSH, appear to be predestined for chat-based LLMs. Compared to conventional high-interaction honeypots for these protocols, using LLMs for honeypots could have the advantage of reducing the risk of unintentional breakouts. Traditional honeypots, designed to replicate real systems, operate within an actual OS environment, which carries the inherent risk that an attacker, upon discovering the honeypot, could exploit vulnerabilities in the underlying OS to escape its confines. Such a breakout would compromise not only the honeypot but also pose a serious threat to the broader network infrastructure. In contrast, LLMs are no real OS; they generate responses based on training data without executing commands in a real system environment. Therefore, an attacker interacting with an LLM-based honeypot can

only engage with the simulated responses generated by the model. Consequently, this approach offers a high level of interaction and realism and ensures a safer and more secure deployment.

The potential for using LLM-based honeypots is even greater when applied to sector- or application-specific protocols, where developing high-interaction honeypots is particularly challenging due to the limited availability of preliminary work. While [2] demonstrated the basic potential of LLMs in simulating various OS environments (Windows, Linux, Mac) and applications (e.g., Jupyter notebooks, TeamViewer), subsequent studies have shown that LLMs can also imitate more specialized environments: MySQL servers [9], IoT devices [10], and even sector-specific devices using protocols such as Modbus and S7comm [11].

Recent research indicates that while code generated by LLMs may appear correct at first glance, it is often prone to errors [12]. Previous studies on LLM-based honeypots have shown that this is also true for this research area, and not all responses generated by LLMs are accurate, suggesting that they are not yet reliable enough to serve as flawless high-interaction honeypots [3,13].

We believe, in the context of honeypots, convincing the attacker takes precedence over achieving flawless accuracy. In realistic scenarios, though, it is essential that the attacker remains unaware of the honeypot's nature for as long as possible. This underscores the importance of evaluating how effectively LLMs can currently function as honeypots and identifying the areas where they still face challenges.

In fact, all prior research approaches to performance evaluation of LLM-based Honeypots are different. Some rely on expert analysis, having humans assess the performance of LLMs [2, 3]. Others create unique metrics, comparing LLM outputs to real server responses at the character or byte level [9, 11], or by contrasting LLM-based honeypots with traditional ones like Cowrie[1].

In the latter case, comparison methods differ, using metrics like successful sessions, average session length, or response success rates [4]. Another approach involves using real server responses as a baseline, comparing the Levenshtein distance and L-ratio between medium-interaction honeypot responses and those generated by LLMs to evaluate performance [13].

Despite their pioneering work, current research on LLM-based honeypots mostly focuses on individual responses without considering the context or the ability to mimic a real system. Although the initial evaluations are valuable, they do not allow for comparisons between different approaches, nor do they clarify whether LLM improvements enhance honeypot effectiveness. As a result, the absence of standardized evaluation metrics limits the ability to assess and validate the true impact of these advancements on honeypots.

At present, there are no established evaluation baselines or ground truth data available for assessing Secure Shell (SSH) sessions, especially those that involve real-world attacks.

With this in mind, we focus on evaluating the performance of GPT-3.5 in mimicking SSH servers as components of honeypots in IT security. Rather than demonstrating its ability to function as a genuine system, our goal is to evaluate how to measure the performance of this fundamental yet state-of-the-art LLM, which serves as the foundation for many rapidly evolving derivative models and represents the first to explore.

To do this, we investigate how well GPT-3.5 can replicate an SSH server, particularly in providing responses that align with attacker requests, maintaining context across multiple interactions, and identifying where the model succeeds or fails. We then explore methods to differentiate between accurate and flawed responses.

In our work, we combine and extend three existing datasets to (1) ensure comparability with prior approaches [3], (2) identify GPT-3.5's limitations with complex single-line commands [14], and (3) maintain realism by including real-world attacks and server responses [15].

As part of advancing the global evaluation of GPT-based honeypots, particularly regarding context-dependent interactions, we contribute:

- An annotation framework[2] for evaluating GPT-based responses as SSH servers, including those involving multiple interactions requiring context, grounded in the Cambridge Dictionary's definition of what is convincing[3].
- Validation of our framework on a subsample of 7,000+ request-response pairs, resulting in 1,400+ unique annotations covering 230 base commands and interaction chains from three state-of-the-art honeypot datasets. GPT-3.5 responses were annotated by five experts, achieving an average Krippendorff's α of 57.4.
- An investigation into GPT-3.5's difficulty in maintaining context over multiple attacker interactions, which can make a honeypot detectable.
- Suggesting that annotation effort and GPT-3.5's performance might be improved through a paraphrase-mining approach, potentially allowing for the distinction between genuine and impostor responses with 77.85% macro F1 using cosine distance.

2 Data Resources

Our first dataset, sourced from the work of [3], involved human participants interacting with an LLM-honeypot system, resulting in 226 unique commands, and is referred to as the Prague dataset, named after TU Prague. Participants used SSH for tasks like package management, file system, and network operations to evaluate their ability to distinguish shelLM outputs from those of a real system.

The second dataset, NL2Bash by [14], was constructed to facilitate the translation of natural language (NL) sentences into Bash commands. It comprises 12,609

[1] github.com/cowrie/cowrie
[2] github.com/TomatenMarc/Dont-Stop-Believin.git
[3] dictionary.cambridge.org/dictionary/english/convincing

text-command pairs initially collected from various web sources such as forums, tutorials, and tech blogs. The dataset includes Bash commands utilizing over 100 unique utilities and corresponding high-quality English descriptions provided by expert Bash programmers. It was used to train and evaluate various neural semantic parsing models, demonstrating the complexity and challenges of mapping NL to Bash commands.

The third dataset, collected to analyze attacker behaviors on a high-interaction Linux SSH honeypot [15], is referred to as the Halle dataset, named after the researchers from Halle University. It was collected over two distinct periods, from May 2017 to September 2019 and January to October 2021. It encompasses the subsequent commands executed by attackers and the responses from an SSH honeypot.

The original datasets can be found or requested from the authors of [3], [14], [15] and are chosen because:

1. **Prague** [3] contains commands used in prior research regarding LLM-based honeypots to have a baseline for comparison.

2. **NL2Bash** [14] includes complex and challenging single-line commands. Responses to those exact commands are usually not directly publicly accessible. The LLM must, therefore, logically capture the meaning of the command and estimate the response.

3. **Halle** [15], captured from real-world attackers interacting with high-interactive SSH honeypots, is the most realistic dataset of these three, containing several coherent commands in a sequence (SSH sessions). This enables us to investigate the influences of contextualization towards the LLM's performance.

For the two datasets without valid server responses (Prague and NL2Bash), we sent all commands to a virtual machine and collected the responses (stdout and stderr). To ensure comparability, we used the same OS (Debian Jessie) as the researchers that created the Halle dataset [15] because responses to the same commands can vary between different Linux distributions. For instance, the `ps` command on Debian might show a different format or include different default columns compared to Fedora or Arch Linux due to differences in how system tools and libraries are configured.

In preparation of the Halle data, we filtered out those commands and responses that contained non-printable characters, e.g., if the attacker used the SFTP subsystem. We then removed duplicates; for example, if an attacker first used the command `id` and then checked system or kernel information via `uname` and then interrupted the connection, we removed all sessions except one. For NL2Bash and Prague, we excluded all commands that produced empty responses from the virtual machine, as these do not meet the criteria for evaluating the usability of metrics via representations. In contrast, empty responses were retained for Halle, as they could be important for the session state, such as when changing the directory directly impacts subsequent commands and responses. Especially for this dataset, every command from each session was provided to the LLM, along with the context of previous commands and responses, ensuring that the LLM could account for the server's state throughout the session.

2.1 Sample Generation

We created a representative sample for each dataset, preserving the original distribution of command length, base commands, and labels. This approach maintained nearly identical mean command lengths and session sizes (for the Halle dataset) while including at least one entry for every base command with a non-empty response. To identify the base commands within the data, we used the method described in [14]. We stripped `sudo` from the beginning of a command and replaced absolute path names with their base names (e.g., `/bin/find` to `find`).

We included commands that are not valid on a command line, such as those in NL2Bash, that remained in plain text rather than being processed into shell commands, assuming a real server to throw errors in such cases. In contrast, the LLM might provide a server response matching the intention [3].

Table 1 provides the differences between the original datasets and the generated sample.

Table 1: **Summary of dataset and sample distributions.** Note that the total count of base commands is not the sum of unique base commands across all datasets, as some base commands appear in multiple datasets.

Dataset	Source	Commands	Unique Base-Commands
Prague	Original	226	67
Prague	Sample	129	49
NL2Bash	Original	12,607	226
NL2Bash	Sample	334	183
Halle	Original	2,483	85
Halle	Sample	943	59
Total	Original	15,316	251
Total	**Sample**	**1,406**	**230**

2.2 LLM Response Generation

We employed *GPT-3.5 Turbo*[4] to simulate an SSH server operating on a Debian Jessie server, chosen specifically because all three datasets utilized this OS, ensuring consistency across our annotations and experiments.

The prompt was meticulously designed to preclude any supplementary commentary or elaboration, thereby guaranteeing that the outputs produced by the language model remained indistinguishable from those generated by an authentic SSH server. To further enhance the realism of the simulation, the prompt explicitly directed the language model to generate contextually appropriate error messages in response to incorrect or malformed commands, effectively emulating the behavior of a genuine server environment.

To leverage the advantages of few-shot learning [8], we provided an example interaction to guide the LLM in understanding the expected format and behavior. This

[4] platform.openai.com/docs/models/gpt-3-5-turbo

example, combined with a carefully structured prompt, was designed to ensure consistency and realism in the LLM's responses throughout the session. The complete prompt, along with the code to retrieve responses from GPT-3.5, our detailed evaluation, sample datasets, and annotation materials, can be accessed in our repository[2].

3 Annotation Framework

As a baseline, five annotators rated the generated responses from the LLM. To ensure consistency in the evaluation process, an annotation guide containing rating instructions and examples of edge cases, carefully deliberated and tested by the main authors, was provided. The annotators were trained students with at least a bachelor's degree in computer science.

The annotators evaluated whether the LLM-generated responses were convincing when compared to real responses from an SSH server. In this context, and according to the Cambridge Dictionary[3], convincing means that the LLM is *able to make you believe that something is true or right*, thereby appearing as if it is a real SSH server.

It is important to note that this does not necessarily equate to technical correctness. A response can be considered convincing even if it contains technical inaccuracies as long as it aligns with the general expectations of how an SSH server might respond under similar circumstances. The focus is on the likelihood of deceiving a potential attacker into thinking they are interacting with a real SSH server.

For this task, a binary annotation system was employed. Annotators rated each response as either convincing or not. Each data point that required annotation consisted of a three-part format designed to simulate an SSH session interaction. This format provided the context for evaluating the LLM's ability to mimic real SSH server responses accurately:

- **Completion Request**: This represents an SSH session history, providing context for the LLM and the annotators. It follows the user-assistant dialogue format, with *user: ¡ssh-command¿* and *assistant: ¡server-response¿*. The session history concludes with a command from a user to which the LLM is supposed to generate a response.

- **Expected Response**: This is the output of the real SSH server in response to the SSH command sequence. It served as a reference for what a typical SSH server might return but was not shown to the LLM. This expected response helped the annotators gauge the accuracy and realism of the LLM's generated response.

- **Completion Response**: This is the LLM-generated response based on the completion request. It is the focus of evaluation. Although the expected response serves as a benchmark, any plausible response that differs from the expected response was still considered valid.

For further details on the annotation, the associated materials, or the results, please refer to our repository[2].

4 Findings

4.1 Annotation Evaluation

In total, the five annotators evaluated 7,030 request-response pairs (1,406 each). Slightly more than half of the pairs were deemed convincing by the annotators, resulting in a fair label distribution. An examination of the individual datasets reveals some differences, as Table 2 shows. The best results were yielded by the Prague dataset with 64% convincing, while the NL2Bash dataset yielded the worst with just 40%.

Table 2: **Results of the annotation the LLM responses as convincing.**

Dataset	Convincing (%)
Prague	64.34
Halle	54.93
NL2Bash	40.12
Overall	52.28

The agreement among the annotators, measured using Krippendorff's strict α, resulted in an average agreement of 57.4 α, which we consider solid given the task's difficulty and the complexity of the data. While the pairs associated with the Prague and NL2Bash datasets showed lower agreement, the Halle dataset achieved a higher agreement of 63.46α, as detailed in Table 3.

Table 3: **Comparison of Inter-Annotator Agreements across the datasets, measured using Krippendorff's α. The overall agreement is calculated as the average performance across the individual datasets.**

	Krippendorff's Alpha (%)
Halle	63.46
NL2Bash	43.90
Prague	40.71
Overall	57.38

4.2 Dataset Comparison

By analyzing the statistical characteristics of the three datasets, we aimed to understand the on-par label distribution. Given that all general parameters, including the prompt and LLM model, were consistent across the three datasets, with the only variations being the commands sent to the LLM and the session history in the Halle dataset, our analysis centered on these specific factors.

First, we analyzed whether certain base commands resulted in a higher level of convincingness than others. The most convincing commands across at least two datasets

with over 20 instances were `whoami`, `crontab`, and `uname`, each rated as convincing in over 90% of cases.

On the other end, we observed base commands mostly performing as non-convincing, such as `w` (4% convincing), `wget` (8% convincing), and `top` (17% convincing).

For `wget` and `top`, the reasons are quite obvious. `wget` is used to download a file from a remote location to the server. The LLM mostly hallucinated the download of that file or stated that the file could not be downloaded for various reasons: not found, DNS not available, etc.

The `top` command retrieves current processes in an interactive environment. The corresponding man page [5] describes that the program provides a dynamic real-time view and requires the use of cursor control keys, such as the arrow keys. Such interactive elements cannot be realized in the current chat-based implementation of GPT-3.5, as it lacks the ability to continuously process and visually display real-time inputs and outputs, as would be needed in an interactive environment. For this reason, attempts to simulate such functions result in error messages. These errors were often rated unbelievable by the annotators (e.g., `top: error while loading shared libraries`).

Similar issues arise with the `w` command, which is meant to display system uptime and information about active users. The LLM had difficulties with this command and mostly stated that it was not found or returned nothing. As it usually returns up-time and logged-in users, including the attacker's connection, the annotators considered this response unconvincing. In contrast to the other unconvincing commands, the reasons here are not obvious. The command is prevalent, non-interactive, and reference responses can be found in several places. One reason could be that the one-letter command is just too short to be unambiguous.

4.3 Command Complexity

Driven by the analysis of individual commands, we now focus on their complexity, as our data indicated that certain commands exhibited varying levels of persuasiveness across different datasets. This variation is particularly evident with the common command `cd`, which was rated as 100% convincing in the Prague dataset, 57% in the Halle dataset, and not convincing at all in the NL2Bash dataset.

To quantify this complexity, we considered the command length and the frequency of certain characters that have a special meaning in a command line and SSH environment [16]. We assume longer commands signal complex operations more often. By using special characters, commands can perform multiple tasks in a single line, handle complex logic, and efficiently manipulate data streams, enhancing both their functionality and perceived complexity.

[5] https://manpages.org/top

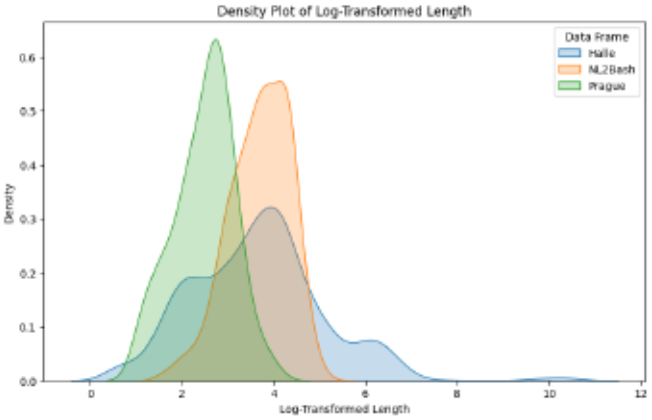

Figure 1: **Density Plot of Log-Transformed Command Length Distribution.**

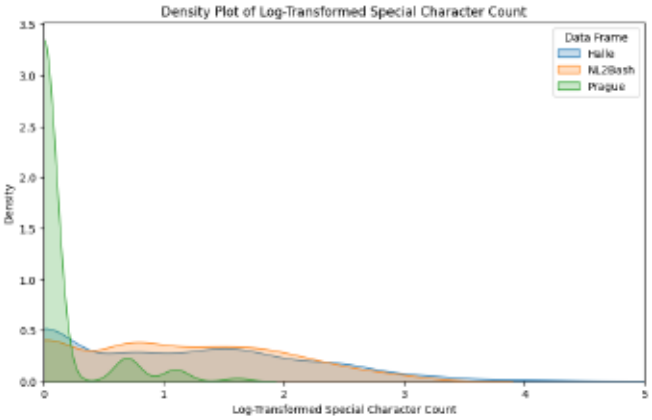

Figure 2: **Density Plot of Log-Transformed Special Character Frequency**

The mean command length varied across the datasets, with Halle having the longest commands at 285.6 characters, followed by NL2Bash at 47.2 characters, and Prague with the shortest commands at 13.4 characters. Focusing on the median reveals that Halle has some large outliers, but the median command length (37 characters) is close to that of NL2Bash (43 characters). Prague's median command length is just slightly below the mean, with 12 characters. However, all command lengths are skewed to the left. When comparing command lengths across datasets, one can observe that Halle and NL2Bash have similar lengths, whereas the Prague dataset primarily consists of shorter commands (Figure 1).

Command length seemingly affects convincingness. Commands longer than the median length exhibit reduced convincingness: 39% for Halle, 32% for NL2Bash. The Prague dataset also shows a slight decrease in convincingness for commands above median length but maintains 59% convincingness. This slightly lower drop in convincingness might be due to the maximum command length of just 56 characters.

The presence of certain characters noticeably reduces the convincingness. Commands containing at least one of the characters ';', '$', '—', '&', or '¡' are only 40% convincing on average in the combined dataset. The presence of

the substitution character '$' reduces the convincingness to 37%, parentheses to 29%, and square brackets to 26%.

Convincingness declines as the number of special characters increases: commands with more than three pipes are convincing in 22% of cases, more than three parentheses in 19%, four or more semicolons or braces in 14%, and four or more angle brackets in 11%.

While the overall presence of special characters generally indicates lower convincingness, no specific group of characters consistently harms the LLM's performance. For instance, in the NL2Bash dataset, the presence of three or more substitution characters '$' results in 8% convincingness, whereas in the Halle dataset, 50% of commands with the same count of substitution characters remain convincing.

We also analyzed the correlation between the use of special characters and convincingness using Kendall's τ. The relationship is weak but statistically significant, with a correlation of -0.21 ($p = 1.509e\text{-}19$) in the overall sample data. Similarly, the correlation between command length and convincingness is also weak but statistically significant, with a correlation of -0.22 ($p = 1.007e\text{-}22$).

Comparing the three datasets, commands in Halle and NL2Bash are more complex than in Prague. The mean number of special characters per command is 6.4 for Halle, 1.5 for NL2Bash, and 0.1 for Prague. Despite some commands in Halle having many special characters, the median in both Halle and NL2Bash is just one, indicating most commands are similar in complexity.

Thus, Halle and NL2Bash are similarly complex in terms of command length and special character count, while the Prague commands are shorter and contain fewer special characters. The distribution of the special characters counts in commands through the datasets can be seen in Figure 2.

4.4 Session State Transitions

To further investigate the differences in convincingness, we examined the role of the session state on the convincingness of the generated responses.

Halle stands out as the most realistic dataset, sourced from real attacker sessions, and is unique in being the only session-based dataset with commands embedded within a session history. For this reason, we focus on Halle, examining transition probabilities to gain deeper insights into its varying scores, specifically analyzing how preceding commands influence the convincingness of subsequent commands within a session.

The data reveals that when a response is not convincing, the likelihood of the subsequent response being believable is only marginally higher (50.96%). However, when a response is convincing, the probability that the following response will also be convincing rises to 59.59%. These findings are detailed in Table 4. The results suggest that while the LLM shows some improvement in generating convincing responses with additional context, there is still a significant risk of inconsistency. This underscores the importance of developing methods to reliably distinguish between genuine and impostor responses, as relying solely on context may not be sufficient to ensure consistently convincing outputs.

Table 4: **Transition probabilities of session state for Halle.**

$p(s+1\|s)$	Non-Convincing	Convincing
Non-Convincing	49.04%	50.96%
Convincing	40.41%	59.59%

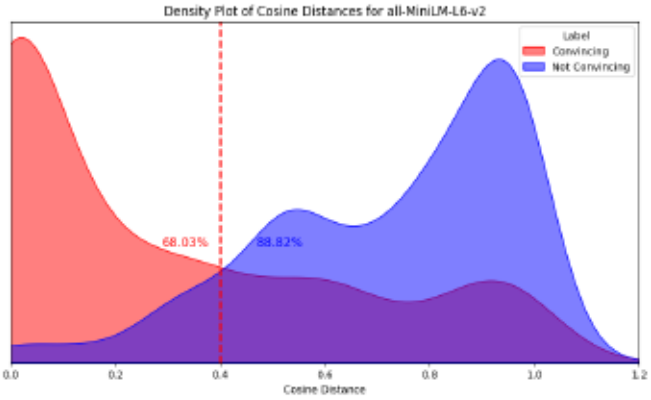

Figure 3: **Density plot of cosine distance illustrating the separation between convincing and non-convincing responses via all-MiniLM-L6-v2.**

4.5 Distance and Representations

To enhance future evaluation of LLM-based honeypots, we conducted a series of ablation experiments with various representation models to establish best practices. By examining the structural differences and similarities between labels, specifically what defines a convincing response versus a non-convincing one in the given context, we aim to gain valuable insights into the measurability of these distinctions.

Sentence-BERT (SBERT) [17] can be used alongside different pre-trained transformer models for comparing text representations. SBERT extends the BERT architecture [18] by utilizing Siamese network structures [19] to produce semantically meaningful sentence embeddings. When these embeddings are combined with cosine distance, a standard method for comparing text representations [20], the classification problem becomes one of metric learning [21], focusing on the measurability of similarities. The classifier leverages pre-trained representations, calculates their distances to their baseline counterparts, and identifies an optimal threshold for effective class separation, thereby streamlining the training.

Cosine distance, $d(v_1, v_2) = 1 - \cos(v_1, v_2) \in [0, 2]$, measures the semantic similarity between two vectors v_1 and v_2 by calculating the cosine of the angle between them ($\cos(v_1, v_2) \in [-1, 1]$) and mapping the result to a range from 0 (similar) to 2 (dissimilar) for better interpretability. When applied to SBERT-generated vector representations, this distance quantifies the semantic similarity between LLM-generated responses (v_1) and baseline outputs (v_2).

We tested various models for classifying the convincingness of LLM outputs compared to the baseline virtual machine outputs using SBERT and cosine similarity.

The evaluation relied on macro F1 scores to assess the models, giving equal importance to both classes, convincing and unconvincing responses. This method ensures that the analysis fairly considers the model's performance across the entire binary classification task without favoring one class over the other.

The models evaluated include all-MiniLM-L6-v2[6], bert-base-uncased[7], distilbert-base-uncased[8], and roberta-base[9]. The macro F1 values for each model are as follows:

- **all-MiniLM-L6-v2:** 77.85%
- **bert-base-uncased:** 75.36%
- **distilbert-base-uncased:** 75.15%
- **roberta-base:** 67.96%

The all-MiniLM-L6-v2 model achieves the highest macro F1 score of 77.85%, outperforming the remaining models with bert-base-uncased and distilbert-base-uncased slightly lower performances. An optimal cosine distance threshold of 0.4 for all-MiniLM-L6-v2 shows that 68.03% of convincing responses have lower distances, while 88.82% of unconvincing ones have higher distances, indicating better semantic similarity for convincing responses, as evidenced in Table 3. All details on all evaluation steps can be viewed in our repository[2].

5 Discussion

We observed differences in the annotation results of the Prague dataset compared to previous reports. Since [3] used a slightly different analysis approach, we first had to align their results with ours. Thereby, we define convincing as the sum of the true-negative (74%) and false-negative (0.5%) rates, considering any response the Prague annotators deemed appropriate as convincing (even if their experts stated it was revealing the honeypot). However, our experts found only 64.3% of the Prague sample to be convincing, compared to the 74.5% in the previous analysis.

Differences in annotation results are common, but the dataset's pre-filtering may have also influenced these outcomes. Removing empty responses might have excluded simpler commands, like creating new folders, changing directories, or deleting files, decreasing perceived convincingness.

To our knowledge, this study includes the largest number of unique base commands used to evaluate such a system, covering the 100 essential command-line utilities[10]. This categorization showed that some base commands yielded higher convincingness while others did not, which we consider a natural phenomenon rather than a flaw. This is consistent with the findings from other studies that have used base commands [3] or command groups [13] for analysis.

However, discrepancies in convincingness between the datasets were observed in the base-command categorization. Listing 1 illustrates the issue[11].

Listing 1: A sample command from the Halle dataset

```
cat /proc/cpuinfo \
| grep name \
| head -n 1 \
| awk '{print $4,$5,$6,$7,$8,$9;}'
```

Although the base command is `cat`, the LLM must also process three other shell commands. Errors from these could mistakenly appear as if they originated from the base command `cat`. Therefore, the base command categorization better describes dataset variation rather than defining the LLM's limitations as a honeypot.

We examined command complexity, finding that the presence of special characters and longer commands correlate with reduced convincingness. The Prague dataset's lower length and complexity may explain its higher scores. However, this does not account for why similar datasets like Halle and NL2Bash do not achieve the same results. The session history might explain this, as transition probabilities indicate that once the LLM generates a convincing response, the likelihood of subsequent convincing responses increases.

Convincingness and annotator agreement fluctuated with more complex datasets. This may be because annotators, like the LLM, struggled with command complexity, and session context aided both. This might explain the higher scores and agreement compared to the NL2Bash and Prague datasets.

However, using SBERT and cosine distance to evaluate GPT-3.5's responses proves effective for future LLM assessments, potentially reducing the need for human annotation.

6 Conclusion & Future Work

In this study, we explored the frontiers of LLM-based honeypots, presenting a novel approach for the performance analysis of such systems. Our findings reveal substantial variability in the performance of GPT-3.5 when employed as a honeypot backend, with convincingness rates across the three datasets ranging from 40% to 64%.

This variability was linked to command complexity; both between and within datasets, longer commands or those with more special characters consistently led to less convincing responses. We highlight GPT-3.5's clear limitations in handling complex SSH commands, rendering it unsuitable as a high-interaction honeypot backend in its current state.

Nevertheless, our study also demonstrates that leveraging the context window of this LLM can significantly

[6] huggingface.co/sentence-transformers/all-MiniLM-L6-v2
[7] huggingface.co/google-bert/bert-base-uncased
[8] huggingface.co/distilbert/distilbert-base-uncased
[9] huggingface.co/FacebookAI/roberta-base
[10] oliverelliott.org/article/computing/ref_unix/
[11] Line breaks inserted for improved readability

enhance overall convincingness, particularly in session-based protocols like SSH, where an initial believable response increases the likelihood of subsequent convincing responses.

Our investigation explored whether assessing the convincingness of generated responses is fundamentally a metric issue. We introduced a new approach using the distance between generated and genuine responses as a measure of believability, with paraphrase mining achieving a macro F1 score of 77.85%. This method shows promise for reducing manual annotation, streamlining LLM performance evaluation, and facilitating comparisons with previous honeypot research. Additionally, it could serve as a pre-filtering tool, automatically rejecting unlikely responses and allowing re-queries, potentially improving overall performance.

Taken together, our work sets the stage for future research comparing LLMs as honeypots, particularly exploring whether fine-tuned models can better handle complex command-line instructions. This is especially relevant for specialized domains like healthcare, where improved interaction fidelity with protocols like HL7 could be highly beneficial. Additionally, our findings suggest further exploration into how these enhancements can keep attackers engaged longer without detecting the honeypot, advancing the effectiveness of LLM-based IT security.

7 Limitations

We identified key limitations in the LLM's practical deployment, particularly with handling time-dependent responses and latency issues. Several commands require the inclusion of real-time data, such as the current date or time, which the LLM lacks awareness of. This gap necessitates integration with external systems or APIs to provide accurate, real-time information. Additionally, while latency was negligible in our theoretical setup, it becomes a significant risk in real-world applications. If the LLM takes too long to respond – especially during complex operations – this delay could tip off an attacker that something unusual is happening. In our study, response generation occasionally took tens of seconds, a delay that would be unacceptable in a real-world scenario where speed is critical. Both issues warrant further investigation, as they are not only relevant for honeypots but also represent broader challenges in LLM research (e.g., [22], [23]).

GPT-3.5's limited context window restricts session history, so we capped Halle dataset sessions at 50 command-response pairs. Longer sessions could be handled through methods like Rotary Position Embeddings and prompt compression [24]. Newer OpenAI models with larger context windows may mitigate this issue, and future studies could benefit from these improved models.

Furthermore, interactive environments challenge LLMs due to limited chat-based interactivity. Solutions include deactivating interactive elements or creating a wrapper to simulate interactions, such as with the vi editor or top.

As research assistants, the annotators carried out the annotation during working hours and were paid appropriately for their time. AI text generation tools, including ChatGPT, were utilized solely for experimental purposes and did not contribute to the writing of this article.

8 Acknowledgments

(Simon B. Weber and Marc Feger contributed equally to this work.) The authors would like to thank their annotators for their expert contributions and the anonymous reviewers for their fair and insightful assessments, which have greatly improved their work. Your willingness to share your expertise and provide constructive feedback is deeply appreciated.

References

[1] M. Abbas-Escribano and H. Debar, "An improved honeypot model for attack detection and analysis," in *Proceedings of the 18th International Conference on Availability, Reliability and Security*, 2023, pp. 1–10.

[2] F. McKee and D. Noever, "Chatbots in a honeypot world," *arXiv preprint arXiv:2301.03771*, 2023.

[3] M. Sladić, V. Valeros, C. Catania, and S. Garcia, "Llm in the shell: Generative honeypots," *arXiv preprint arXiv:2309.00155*, 2023.

[4] Z. Wang, J. You, H. Wang, T. Yuan, S. Lv, Y. Wang, and L. Sun, "Honeygpt: Breaking the trilemma in terminal honeypots with large language model," *arXiv preprint arXiv:2406.01882*, 2024.

[5] P. P. Ray, "Chatgpt: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope," *Internet of Things and Cyber-Physical Systems*, vol. 3, pp. 121–154, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S266734522300024X

[6] S. Biswas, "Role of chatgpt in computer programming." vol. 2023, p. 9–15, Feb. 2023. [Online]. Available: https://mesopotamian.press/journals/index.php/cs/article/view/51

[7] M. Zhang and J. Li, "A commentary of gpt-3 in mit technology review 2021," *Fundamental Research*, vol. 1, no. 6, pp. 831–833, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2667325821002193

[8] "Language models are few-shot learners," in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., vol. 33. Curran Associates, Inc., 2020, pp. 1877–1901. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf

[9] Y. Hu, S. Cheng, Y. Ma, S. Chen, F. Xiao, and Q. Zheng, "Mysql-pot: A llm-based honeypot for mysql threat protection," in *2024 9th International*

Conference on Big Data Analytics (ICBDA). IEEE, 2024, pp. 227–232.

[10] V. S. Mfogo, A. Zemkoho, L. Njilla, M. Nkenlifack, and C. Kamhoua, "Aiipot: Adaptive intelligent-interaction honeypot for iot devices," in *2023 IEEE 34th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. IEEE, 2023, pp. 1–6.

[11] C. Vasilatos, D. J. Mahboobeh, H. Lamri, M. Alam, and M. Maniatakos, "Llmpot: Automated llm-based industrial protocol and physical process emulation for ics honeypots," *arXiv preprint arXiv:2405.05999*, 2024.

[12] J. Liu, C. S. Xia, Y. Wang, and L. ZHANG, "Is your code generated by chatgpt really correct? rigorous evaluation of large language models for code generation," in *Advances in Neural Information Processing Systems*, A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, Eds., vol. 36. Curran Associates, Inc., 2023, pp. 21 558–21 572. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2023/file/43e9d647ccd3e4b7b5baab53f0368686-Paper-Conference.pdf

[13] J. Ragsdale and R. V. Boppana, "On designing low-risk honeypots using generative pre-trained transformer models with curated inputs," *IEEE Access*, vol. 11, pp. 117 528–117 545, 2023.

[14] X. V. Lin, C. Wang, L. Zettlemoyer, and M. D. Ernst, "Nl2bash: A corpus and semantic parser for natural language interface to the linux operating system," *arXiv preprint arXiv:1802.08979*, 2018.

[15] M. Knöchel and S. Wefel, "Analysing attackers and intrusions on a high-interaction honeypot system," in *2022 27th Asia Pacific Conference on Communications (APCC)*. IEEE, 2022, pp. 433–438.

[16] M. Schröder and J. Cito, "An empirical investigation of command-line customization," *Empirical Software Engineering*, vol. 27, no. 2, p. 30, 2022.

[17] N. Reimers and I. Gurevych, "Sentence-bert: Sentence embeddings using siamese bert-networks," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 11 2019. [Online]. Available: https://arxiv.org/abs/1908.10084

[18] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, J. Burstein, C. Doran, and T. Solorio, Eds. Minneapolis, Minnesota: Association for Computational Linguistics, Jun. 2019, pp. 4171–4186. [Online]. Available: https://aclanthology.org/N19-1423

[19] J. Bromley, I. Guyon, Y. LeCun, E. Säckinger, and R. Shah, "Signature verification using a "siamese" time delay neural network," in *Proceedings of the 6th International Conference on Neural Information Processing Systems*, ser. NIPS'93. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1993, p. 737–744.

[20] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in Neural Information Processing Systems*, C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Weinberger, Eds., vol. 26. Curran Associates, Inc., 2013. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf

[21] S. Chopra, R. Hadsell, and Y. LeCun, "Learning a similarity metric discriminatively, with application to face verification," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, 2005, pp. 539–546 vol. 1.

[22] J. Xu, R. Zhang, C. Guo, W. Hu, Z. Liu, F. Wu, Y. Feng, S. Sun, C. Shao, Y. Guo *et al.*, "vtensor: Flexible virtual tensor management for efficient llm serving," *arXiv preprint arXiv:2407.15309*, 2024.

[23] X. Ning, Z. Lin, Z. Zhou, H. Yang, and Y. Wang, "Skeleton-of-thought: Large language models can do parallel decoding," *arXiv preprint arXiv:2307.15337*, 2023.

[24] H. Jiang, Q. Wu, X. Luo, D. Li, C.-Y. Lin, Y. Yang, and L. Qiu, "Longllmlingua: Accelerating and enhancing llms in long context scenarios via prompt compression," *arXiv preprint arXiv:2310.06839*, 2023.

Chapter 6

A Novel Approach to Medical Device IT Security Landscape Analysis Leveraging Manufacturer Disclosure Statements

This chapter gives an overview of the contributions and the impact of the paper Stein et al. (2024)[1]:

Stefan Stein, Simon Weber, Michael Pilgermann, Thomas Schrader

"A Novel Approach to Medical Device IT Security Landscape Analysis Leveraging Manufacturer Disclosure Statements"

In: *IEEE Access*, Volume 12, pages 160506 - 160515, Oktober 2024
Acceptance Rate: ~27%

6.1 Summary

This paper presents a systematic analysis of MDS2 as a framework for assessing the IT security landscape of medical devices. Traditional methods for evaluating medical device security often rely on high-level data sources that lack deeper insights. This study reviews 147 MDS2 documents across different versions (2008, 2013, and 2019), assessing their suitability in depicting a comprehensive security posture.

The authors find that MDS2 documents provide a structured approach to evaluate device security attributes, including encryption, access control, and logging capabilities. The analysis shows that MDS2 documents allow intra-version comparability for the versions 2013 onward and inter-version evaluations, although limitations, such as question phrasing variations, limit direct cross-version comparisons. The study identifies significant gaps, such as restricted malware protection flexibility. It also notes inconsistencies in MDS2 documents, posing challenges

[1]

for automated analysis. To enhance MDS2's effectiveness, the authors formulate several recommendations for improvement of future MDS2 versions, e.g., expanding the existing sections regarding SBOM to include hardware and operational elements, and adopting frameworks like the Common Platform Enumeration (CPE) to streamline vulnerability management. Furthermore, the authors argue a centralized MDS2 repository for up-to-date, machine-readable access would enable deeper insights into the evolving security landscape of medical devices, supporting healthcare providers and researchers.

6.2 Personal Contribution

Stefan Stein had the idea for this study and designed the study framework. The research questions were collaboratively formulated by Simon Weber and Stefan Stein. Stefan Stein conducted the preparation and processing of the MDS2 documents. The findings were derived, structured, and intensively discussed by Stefan Stein and Simon Weber. The manuscript was written collaboratively by Simon Weber and Stefan Stein, with feedback on drafts provided by Michael Pilgermann, Thomas Schrader, and Martin Sedlmayr.

6.3 Importance and Impact on this Thesis

This paper introduces a long-needed solution to answer open questions that operators and researchers face regarding risk assessment and decision-making in medical device security. The cross-evaluation of MDS2 documents provides a unique advantage in answering these questions more comprehensively than previous methods used for security landscape analysis, as these documents are specifically tailored to the security requirements of medical technology. Additionally, the historical availability of MDS2 documents allows for the identification of IT security trends over time, offering insights that can guide future research endeavors. This paper highlights immediate opportunities to address pressing questions for operators and lays the groundwork for researchers to shape the future of medical device security.

A Novel Approach to Medical Device IT Security Landscape Analysis Leveraging Manufacturer Disclosure Statements

STEFAN STEIN
University of Applied Sciences
Brandenburg a. d. Havel, Germany

SIMON B. WEBER
Heinrich-Heine-University
Düsseldorf, Germany

MICHAEL PILGERMANN
University of Applied Sciences
Brandenburg a. d. Havel, Germany

THOMAS SCHRADER
University of Applied Sciences
Brandenburg a. d. Havel, Germany

MARTIN SEDLMAYR
University of Technology
Dresden, Germany

Abstract

The growing number of cyberattacks targeting the healthcare sector increasingly threatens network-enabled medical devices that are vital for life-sustaining patient care. Security researchers and healthcare IT managers are pursuing effective methods to assess the IT security landscape of medical devices. Their goal is to develop a comprehensive understanding of the devices' IT security status. Recent studies have successfully uncovered structural deficiencies in medical device security. However, the limitations of their data sources, particularly in evaluating features like logging capabilities and third-party libraries, restrict the scope of their findings. In this study, we present the first systematic analysis of Manufacturer Disclosure Statement for Medical Device Security (MDS2) documents to evaluate their use in creating holistic statements regarding the IT security posture of medical devices. We examined a total of 147 MDS2 documents encompassing devices from 105 different classes. Our findings indicate that MDS2 documents, especially those from the second version (2013) onwards, are suitable for this purpose. We also discuss the shortcomings of the latest MDS2 version in meeting current IT security requirements. Based on the gaps identified, we developed several recommendations to improve MDS2 documents and enhance their effectiveness across the global healthcare sector. In the future, these documents could be used not only for comprehensive landscape analyses but also for organization-specific reports, providing healthcare managers with direct insights into the IT security status within their institutions.

Index Terms

Cyber Security, IT Security, Manufacturer Disclosure Statement for Medical Device Security, MDS2, Medical Cyber-Physical Systems, MCPS, Medical Device, Medical Information System, OT Security, Vulnerabilities

1 Introduction

Between 2016 and 2023, more than 4,700 medical institutions worldwide were victims of cyberattacks [1]. The latest situation report from the German Federal Office for Information Security (BSI) in 2023 highlights the severity of this issue, with the healthcare sector in Germany being the most affected, reporting 132 out of 490 incidents in critical infrastructure [2]. Reports from security authorities, such as the US Department of Health and Human Services (HHS) and the European Union Agency for Cybersecurity (ENISA), provide insights into the sector's security challenges.

Given the critical nature of medical devices in healthcare, their vulnerabilities pose a unique challenge. These devices are repeatedly cited as key factors in cyber threats [3–5]. However, these reports tend to provide only a high-level view of the security deficiencies in medical devices, lacking detailed insights.

In the pursuit of a detailed but comprehensive IT security assessment of medical devices, researchers have employed various data sources. [6] utilized product summaries from the Food and Drug Administration (FDA) to derive the needed IT security features. While an important step, the authors acknowledged that these summaries do not require extensive IT security information, leaving potential gaps in the data. [7] leveraged open contracting data from national health services across 36 countries to generate medical device asset lists, correlating these lists with open-source IT security databases, such as Common Vulnerabilities and Exposures (CVE), National Vulnerability Database (NVD), and ICS Medical Advisory (ICSMA). Similarly, [8] analyzed vulnerabilities from the NVD, identifying electronic health records, wireless infusion pumps, endoscopic cameras, and radiology information systems as exceptionally vulnerable. Although these studies successfully uncover structural deficiencies in medical device security, their reliance on vulnerability

data limits their scope, particularly in evaluating features like logging capabilities or utilized third party libraries, which are not captured in typical vulnerability databases.

The Manufacturer Disclosure Statement for Medical Device Security (MDS2) was developed to provide a standardized and structured method for medical device manufacturers to disclose the security features of their products. These documents serve as a tool for healthcare organizations to assess the IT security of medical devices they use. Initially designed as a questionnaire with 43 questions to help device operators evaluate key security aspects, MDS2 has evolved into a comprehensive framework. The latest version, released in 2019, includes 216 questions across 23 security-related categories, covering aspects such as encryption, access control, and logging capabilities. An important element of MDS2 is the integration of the Universal Medical Device Nomenclature System (UMDNS), which standardizes device classification and facilitates the comparison of security information across different products. The UMDNS contains approximately 5,000 major categories for all items used in the medical field. The widespread adoption of MDS2 documents, particularly in the U.S. and increasingly in international settings [9], presents a valuable opportunity for deriving a comprehensive security landscape of medical devices.

Compared to the previous approaches, the derivation of security capabilities from MDS2 documents has the following advantages:

1. Comprehensive coverage of security features beyond just vulnerable components.
2. Information specifically tailored to medical devices.
3. Increasing international adoption, enhancing relevance.

We contribute to the global assessment of IT security characteristics of medical devices:

- Conducting the first analysis, to our knowledge, of how MDS2 documents can be used to compare IT security features across different medical devices and MDS2 versions.
- Investigating the potential of MDS2 documents, beyond their primary purpose, to assess the broader security landscape of medical devices.
- Evaluating the integration of current security standards and frameworks in the latest version of MDS2 and proposing enhancements to even better reflect the IT security of medical devices.
- Offering suggestions to improve the usability of MDS2 documents for operators and researchers based on the framework conditions for their distribution and usage.

2 Methodology

In this study, four research questions were developed to examine both the technical and content-related requirements necessary for a comprehensive evaluation of MDS2 documents. The research aims to derive conclusions regarding the utility and efficacy of these documents to explore the landscape of medical device IT security.

2.1 Research Questions

RQ 1 - Are structure and format of MDS2 documents suitable for enabling automated analysis and deriving insights into the IT security posture of medical devices?

According to the Healthcare Information and Management Systems Society (HIMSS) and the National Electrical Manufacturers Association (NEMA), one of the purposes of MDS2 documents is to enable a scalable IT security assessment while maintaining compliance and structure. The published content of the documents should ideally meet the standard's requirements for machine-readability, information provision, updates, and complete answers to all questions in each category. The documents' formats should differ only slightly in structure and processing options so that a configured parser can extract all the necessary information from them with the same quality. We analyze whether the documents are available in the defined structure, whether specifications are compliant with standards, and how data extraction can be performed. In addition, metadata such as creation date, creation software, content, document size, document length, and graphical implementation are considered.

RQ 2 - Can MDS2 documents be used to reliably compare the security-related attributes of medical devices intra-version?

In order to enable a reliable comparison of security-related attributes of medical devices, it is essential that the questions and manufacturers' responses in MDS2 documents go beyond meeting formal and structural requirements. They must contain detailed and relevant information that allows for the evaluation and comparison of IT security levels. One of the core objectives of MDS2 is to provide information on the medical device's physical, technical, and administrative functions and the manufacturer's recommended security measures. To achieve comparability, it is crucial that the variability in the way manufacturers respond to the same question remains minimal. Previous research has identified specific areas that serve as reliable indicators of the overall security posture of medical devices [10–12]. These include anti-malware software, password regulation policies, the operating system utilized, disclosure of used third-party software, hardening measures, and provided security documentation for operators. For this study, these categories were used as the basis for selecting relevant questions in the MDS2 documents that allow for meaningful comparisons of security attributes.

RQ 3 - Can MDS2 documents be used to reliably compare the security-related attributes of medical devices inter-version?

The analysis of MDS2 documents across versions presents an opportunity to analyze the evolution of IT security

levels. By examining how manufacturers' responses to specific security-related questions evolve over time, it is possible to assess whether improvements or deteriorations in IT security measures have occurred. Measuring these inter-version differences allows for an informed understanding of the general trend in security advancements or potential gaps.

RQ 4 - How might future MDS2 versions be revised to increase their value for medical device operators and facilitate security landscape analyses?

The objective of this research question is to examine whether the most recent MDS2 version adequately integrates current IT security frameworks, standards, and technological guidelines. Relevant standards include recent publications from the National Institute of Standards and Technology (NIST), ENISA, the Cybersecurity and Infrastructure Security Agency (CISA), and other governmental agencies. Consideration of current standards such as the NIST Special Publication (SP) 800 series, ISO 27000, and IEC 62443 can help manufacturers and operators to secure medical devices. The goal is to develop suggestions that could be incorporated into future MDS2 versions to meet the latest IT security requirements, improve IT security of medical devices, and ensure future readiness for IT security landscape analyses.

2.2 Evaluation procedure

Given the limited availability of MDS2 documents online, this study's sources included medical institutions, hospitals, federal authorities, associations, commercial companies, and manufacturers. The evaluation started with a metadata review and visual inspection to assess the documents' structure, content, and appearance. A subsequent structural analysis ensured completeness, proper sequencing, and the inclusion of essential cross-references and appendices, which are vital for usability and seamless system integration. The evaluation also investigated the potential for automating the analysis of these documents to assess their suitability in deriving a security landscape for medical devices. Due to the heterogeneous nature of the file formats and the diversity in visual representations, the data extraction methods were adapted based on format and readability. Machine-readable formats enabled direct text extraction, while Optical Character Recognition (OCR) technology was applied to non-readable PDFs. In cases where both methods were inadequate, manual extraction was employed as a fallback strategy. In order to detect and remove duplicate entries, an MD5 hash value was generated for each of the MDS2 documents.

A detailed content analysis focused on evaluating the coherence and depth of responses, especially in recent versions of the documents, where open-ended and more elaborate answers are increasingly common. For instance, the 2019 version introduced requests for additional information beyond basic Yes, No, or N/A responses. The evaluation paid particular attention to the relevance and specificity of these responses in addressing security requirements. Furthermore, key questions were identified that would facilitate cross-version comparisons, enabling a more nuanced and in-depth analysis in subsequent stages of the research.

3 Results

A total of 147 MDS2 documents were collected from various sources. These documents contain detailed descriptions of the IT security characteristics of medical devices produced by 48 different manufacturers. To estimate the range of device classes represented in this dataset, UMDNS data, as provided by the manufacturers, was used. In total, the documents describe 105 different classes of medical devices. However, due to the absence of publicly available information regarding the total number of MDS2 documents for all medical devices, no definitive conclusions can be drawn about the representativeness of the sample. Additionally, it is not possible to precisely determine the number of networkable medical device groups. The UMDNS classification includes both networkable and non-networkable devices, such as waste containers and scalpels, which complicates efforts to isolate only networkable groups within the dataset.

RQ 1 - Are structure and format of MDS2 documents suitable for enabling automated analysis and deriving insights into the IT security posture of medical devices?

To determine whether MDS2 documents are suitable for automated analysis, we examined the three available versions: 2008 (38 documents), 2016 (56 documents), and 2019 (53 documents). Table 1 presents the characteristics of these document versions, highlighting the variations in their size and scope both within and between the versions. The 2008 MDS2 documents are primarily in non-machine-readable PDF formats, necessitating conversion for automated processing. This conversion is further complicated by the fact that some manufacturers decided to complete the documents by hand. The 2013 version remains largely in PDF format, but structural improvements made the use of OCR technology feasible. Nevertheless, 15% of the 2013 documents deviated considerably and required manual adjustment efforts.

Table 1: **Statistics on the MDS2 documents examined by version**

	Number of pages			Number of questions	Number of categories
	Avg	Median	Max		
2008	2,8	2	12	41	4
2013	8,9	7	76	85	20
2019	21,9	15	109	216	23

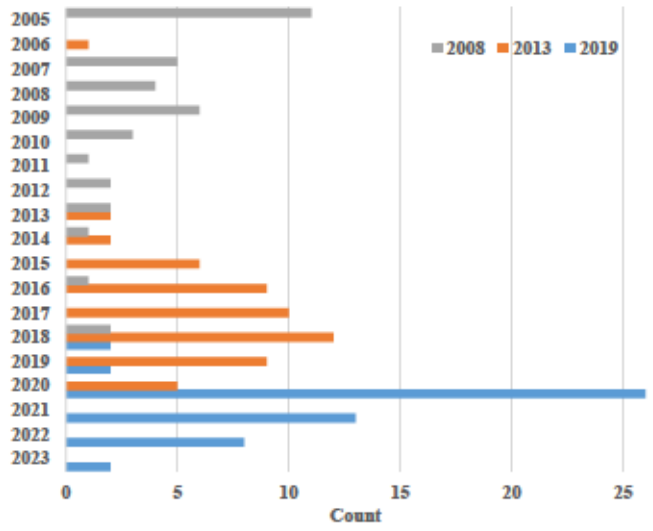

Figure 1: **Overview of the absolute publications of MDS2 documents per year in relation to the versions.** Note that, according to metadata, some documents were created before the MDS2 version used was published.

The 2019 MDS2 documents, while the best machine-processable of the three versions, still exhibit considerable variability. Manufacturers have increasingly used these documents to provide additional information beyond the intended scope, such as cross-references to standards like IEC TR 80001 [13], NIST SP 800-53 [14], and ISO 27002 [15]. These additions, not envisaged by the MDS2 standard, often include extra columns and custom formatting, introducing further complexities for automated data extraction. Additionally, formatting errors, such as incorrect abbreviations and missing questions, were observed in this version. These structural differences and formatting errors can complicate automated data extraction, but appropriate pre-processing and cleaning steps can minimize these challenges. After manual corrections, the 2013 and 2019 MDS2 documents were converted to machine-readable formats, making them generally suitable for automated IT security assessments. Modern technologies like OCR or natural language processing (NLP) enhance this process.

Inconsistencies in metadata, particularly regarding the documents' release dates, were discovered during our analysis. Across all three MDS2 versions, we identified documents with metadata indicating publication before the official release of the standard. For example, some 2008 documents were allegedly published in 2005 and 2007, while metadata on a 2013 document suggested its creation in 2006. Similarly, some 2019 documents were dated 2018. Figure 1 visualizes these discrepancies. This anomaly may be attributed to manufacturers updating MDS2 documents for previously released devices in response to a new MDS2 version and backdating them to align with previous timelines.

The publication trends underscore the growing market acceptance of the MDS2 standard. Notably, the time required to reach peak publication rates has steadily decreased across the different versions. While the 2013 version required approximately 4-5 years to achieve widespread adoption, the 2019 version reached its peak publication rate in just one year. This accelerated dissemination highlights the increasing relevance and integration of the MDS2 standard in the industry, reflecting its growing importance in medical device security practices.

RQ 2 - Can MDS2 documents be used to reliably compare the security-related attributes of medical devices intra-version?

Results for MDS2 2008 version

The 2008 version of the MDS2 documents is not suitable for reliable intra-version comparison of security-related attributes in medical devices. The primary issue lies in the overly narrow scope of the questions, which limits the depth and breadth of the information that manufacturers can provide. This issue is particularly evident in matters concerning electronic protected health information (ePHI), where the questions are focused on very specific aspects of device functionality, often omitting broader security considerations. A clear example of this restrictive focus is the scope of questions related to network security. Instead of addressing broader network capabilities of medical devices, questions focus on whether ePHI is transmitted over a network (question 4e) or via a wireless connection (question 4f). Similarly, question 18 asks if network connections that handle ePHI are encrypted, leaving out other data transmissions. Consequently, manufacturers could answer positively to security-related questions even if those security measures are limited to specific aspects of the device's functionality. For example, if a device allows remote access without encryption but does not handle ePHI over that connection, the manufacturer could answer "no" to the questions about network transmissions without reflecting on the broader security risk posed by unencrypted access. Additionally, a manufacturer could affirmatively answer the question on encrypted ePHI transmission, even if unencrypted remote access exists, provided that ePHI is transmitted over a separate encrypted channel. The logging capabilities exhibit a similar ePHI-centric approach. Question 15 focuses heavily on ePHI-related events, with three of its four subquestions concerning actions involving ePHI: (b) viewing, (c) creation/modification/deletion, and (d) import/export or transmittal/receipt of ePHI. This narrow scope means that a manufacturer could claim comprehensive logging without accounting for other critical security events, such as access control violations, security alerts, or blocked data streams.

Beyond the strong focus on ePHI, the questions are partly framed so narrowly that a "no" response may paradoxically indicate a higher security standard than a "yes" response. For example, question 13 asks, "Does the device support user/operator specific ID *AND* password?" A manufacturer using fixed user IDs (e.g., admin, operator) combined with a robust password policy would be forced to answer "no" due to the restrictive nature of the question. In contrast, a manufacturer allowing customization of both ID and password, but without enforcing strong password policies, could answer "yes."

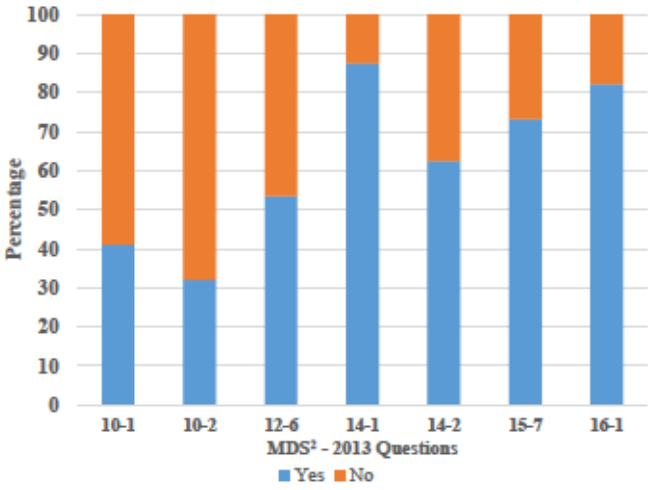

Figure 2: **Analysis of selected questions from the 2013 version:**
[10-1] Does the device support the use of anti-malware software?
[10-2] Can the device owner install or update anti-virus software?
[12-6] Can the device be configured to enforce creation of user account passwords that meet established complexity rules?
[14-1] In the notes section, list the provided or required (separately purchased and/or delivered) operating system(s).
[14-2] Is a list of other third-party applications provided by the manufacturer available?
[15-7] Are all communication ports which are not required for the intended use of the device closed/disabled?
[16-1] Are security-related features documented for the device user?

In this case, the affirmative response indicates a weaker security posture, as it may not meet modern security requirements for complex, unique, and long passwords [16].

For these reasons, the security level of different medical devices cannot be reliably compared based on their MDS2 2008 documents. The actual security posture of a device cannot be accurately derived from the responses, as the narrowly framed questions fail to capture the broader security context and critical attributes beyond the specific scenarios they address.

Results for MDS2 2013 version

The documents in the 2013 version were largely comparable within their version. As explained in section 2, questions about malware detection, authentication, operating system and third-party components, hardening measures, and security documentation were used for the intra-version comparison of security features: Figure 2 provides the distribution of responses for these questions. 88% of the documents provided information about the operating systems used (14-1) whereas only 62% included details regarding third-party software used (14-2). Moreover, the data indicated that 68% of the devices associated with MDS2-2013 documents did not permit users to install additional security software (10-2).

When asked about password complexity, 46% of respondents to question 12-6 reported that it is not possible to create or enforce complex passwords on the device. Additionally, 73% of manufacturers disabled unnecessary communication ports (15-7), and over 80% provided security-related documentation for operators (16-1).

In comparison to the 2008 version, the 2013 MDS2 version introduced revised questions and reduced emphasis on ePHI. The 2013 version addressed all types of network connections, including those involving ePHI, when evaluating network capability. This broadened scope enables a comparison of IT security features across different medical devices based on MDS2-2013 documents. Furthermore, the 2013 version demonstrated a high level of completeness and consistency in the answers provided, with 99% of the responses following the format of "Yes," "No," "N/A," or "See Notes." A small number of responses were marked with a line, which was interpreted as "No." The structure of the notes section presented challenges during the analysis, as 22% of the documents contained repeated notes, failed to adhere to format requirements, or lacked notes in the appendix entirely. Beyond these inconsistencies, the 2013 documents included very few cross-references or appendices, which improved the overall comparability of the answers.

Results for MDS2 2019 version

The 2019 MDS2 documents were suitable for intra-version comparisons, as they exhibited a largely consistent structure and provided more detailed documentation of security measures compared to previous versions. This consistency allows for effective comparison of IT security aspects across various medical devices. Furthermore, the 2019 MDS2 documents demonstrate an improvement in informativeness, with the number of questions increasing by a factor of 2.5 compared to the previous version. This considerable expansion provides a more comprehensive basis for security evaluation and comparison. Specific questions from each aforementioned security domain were selected for intra-version comparison within the 2019 MDS2 documents.

The evaluation of security characteristics (Figure 3) across medical devices using the 2019 MDS2 documents shows several notable trends. Approximately 81% of the devices reported the use of an operating system (CSUP-2) and 72% of devices indicated the availability of a Software Bill of Materials (SBOM), which is crucial for identifying and managing security risks associated with third-party components. Yet only 28% of the 2019 MDS2 documents directly included an SBOM list. In terms of malware protection, 66% of the responses indicated that anti-malware software was installed on the devices (MLDP-2), although only 28% of the documents stated that anti-malware software could be either installed or updated by the operator (MLDP-2.3). The logging of antivirus messages was noted in nearly two-thirds of the MDS2 documents, but only 15% of those systems displayed such messages in the user interface, highlighting a gap between system logging capabilities and user-facing notifications.

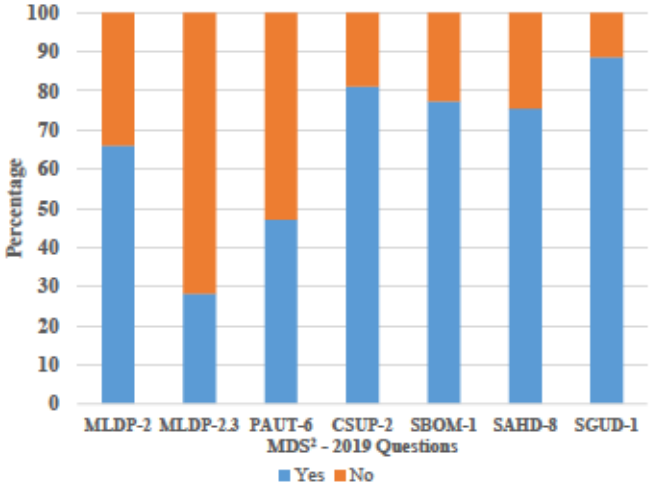

Figure 3: **Analysis of selected questions from the 2019 version:**
[MLDP-2] Does the device contain Anti-Malware Software?
[MLDP-2.3] Does the device documentation allow the owner/operator to install or update anti-malware software?
[PAUT-6] Is the device configurable to enforce creation of user account passwords that meet established (organization specific) complexity rules?
[CSUP-2] Does the device contain an Operating System?
[SBOM-1] Is the SBoM for this product available?
[SAHD-8] Are all communication ports and protocols that are not required for the intended use of the device disabled?
[SGUD-1] Does the device include security documentation for the owner/operator?

Moreover, 47% of the documents indicated that there was no possibility for enforcing a custom password policy (PAUT-6). 75% of manufacturers reported that unnecessary communication ports were disabled by default (SAHD-8), demonstrating proactive measures to reduce attack surfaces. Furthermore, 89% of the documents stated that security documentation is provided to users (SGUD-1), reflecting a strong trend towards better guidance on device security management.

RQ 3 - Can MDS2 documents be used to reliably compare the security-related attributes of medical devices inter-version?

Due to the narrowly formulated questions, it is impossible to compare the IT security capabilities of medical devices based on single questions from MDS2 documents of the 2008 version to other versions. Comparisons between the 2013 and 2019 documents based on similar questions also provide limited cross-version insights, as the phrasing of questions changed over time, which could lead to different responses. For instance, while the 2013 version requires the device to be configurable to enforce any preconfigured password policy (question 12-6), the 2019 version asks whether the device allows the configuration of custom, organization-specific password policies (question PAUT-6).

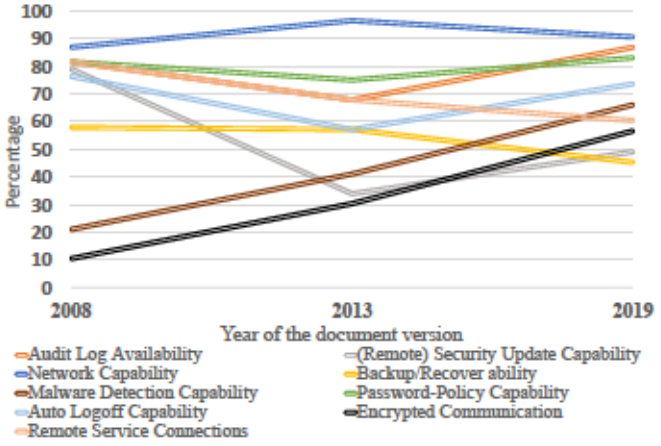

Figure 4: The coverage of security-relevant categories per MDS2-version We attribute the apparent deterioration from 2008 to 2013 documents to changes in the MDS2 version.

This shift in emphasis from enforcing pre-existing rules to supporting customizable policies could result in different responses, even though both questions address the enforcement of password complexity. Such differences in requirements could contribute to the slightly lower fulfillment rate in the 2019 version (54% vs. 47%), but not necessarily indicate a deterioration of the security level.

Thus, instead of comparing responses to specific questions across versions, a category-based approach was adopted. The categories were formulated based on insights from Section 2.1, a survey of analyzed attack types [17], and a review of healthcare data breach causes [18]. These findings led to the identification and selection of the following categories for analysis across all document versions: *audit logs, backups, automatic logoff, security updates, malware/AV detection, encrypted communications, network compatibility, password policies*, and *remote serviceability.*

The results of the analysis are shown in figure 4. Devices with a 2008 MDS2 document exhibited 87% network compatibility, which increased to over 90% for devices covered by later versions. This underscores that MDS2 documents are especially useful for analyzing the security characteristics of network-capable medical devices.

Six categories exhibited stagnation or decline from the 2008 to the 2013 document version, but four of these categories show a reversal by 2019, with three reaching or surpassing 2008 levels. The category *(remote) security update capability*, in contrast, fell sharply from around 80% in the 2008 document version to approximately 35% in the 2013 version before recovering slightly to just under 50% in the 2019 version. The related category of remote service connection availability followed a downward trend, starting at 82% in the 2008 documents, dropping to 68%, and then declining further to 60% in the 2019 version. The backup-recover capability category also exhibited a downward trend, stagnating at 58% between 2008 and 2013 versions, before falling to 45% in 2019 version documents. Two categories demonstrated continuous improvement: *malware detection* and *encrypted communication* increased with each version, suggesting an overall enhancement of protective measures over time.

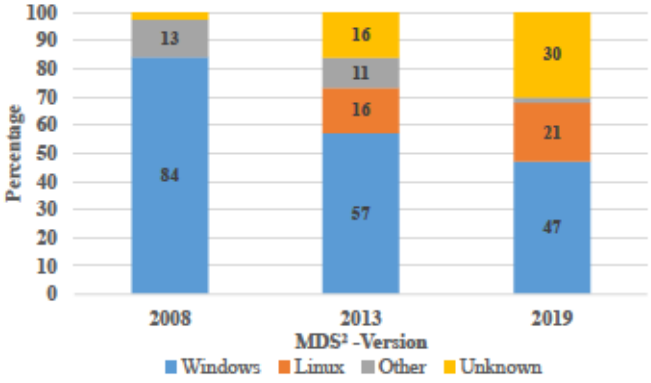

Figure 5: **The distribution of operating systems.** In 2008, the most used operating system was Windows, which gradually changed in 2013 and 2019 to a more heterogeneous distribution in systems.

The evaluation of the operating systems used (figure 5) reveals a high prevalence of Windows operating systems, with over 80% adoption in the 2008 version. From version 2013 onwards, manufacturers began to incorporate more Linux derivatives. By the 2019 version, while Windows operating systems still account for nearly 50%, there is a visible trend towards their replacement by Linux derivatives.

The security level of medical devices is closely tied to the operating system installed, as each system presents different levels of exposure to vulnerabilities. For instance, in an overview of known vulnerabilities in 2013, Microsoft ranked third with 344 reported vulnerabilities in the operating system, software, and hardware areas, while Red Hat, the developer of Red Hat Enterprise Linux (used by 12.5% of Linux-based devices), ranked ninth with 131 vulnerabilities in the same areas [19]. According to data from the NIST database, the entire family of Windows operating systems, starting in 1999, has 7387 known vulnerabilities [20]. In comparison, all Linux derivatives combined have 2,353 known vulnerabilities until 30th August 2024 [21]. These findings suggest that the choice of operating system impacts the security posture of medical devices. Devices running operating systems with a higher number of vulnerabilities may require additional security hardening or protective measures to mitigate the risk of exploitation. Therefore, the increasing adoption of operating systems with verifiably fewer vulnerabilities, such as Linux derivatives, could indicate a trend towards enhanced security in medical devices.

RQ 4 - How might future MDS2 versions be revised to increase their value for medical device operators and facilitate security landscape analyses?

Government agencies, including NIST, CISA, and the Internet Engineering Task Force (IETF), provide various frameworks and regulations to improve the security of information systems. For example, NIST SP 800-128 [22] and ISO 27001 [23] advocate the integration of a Security Information and Event Management (SIEM) system, while the FDA [24] recommends this approach in cases where devices present cybersecurity risks.

The 2019 MDS2 version already addresses SIEM-related functionality by including the question of whether audit log content from a device can be exported to a SIEM solution. However, future MDS2 revisions could further expand on this functionality by incorporating SIEM rules in SIGMA format. This would allow for immediate implementation of network anomaly detection, potentially enhancing IT security even for legacy medical devices. Moreover, the inclusion of a Security Orchestration, Automation, and Response (SOAR) component, as recommended by NIST SP 800-92 [25], could offer higher levels of automation and scalability, allowing for future-proof security operations and policies. Integrating SOAR with SIEM systems optimizes the use of real-time data and increases the overall security posture of medical device ecosystems.

Supply chain security is another important area for future MDS2 revisions. NIST SP 800-161 [26] and Executive Order 14028 [27] highlight the importance of SBOMs for ensuring supply chain transparency. The Medical Device Cybersecurity Working Group advocates that SIEM systems should incorporate SBOM processing capabilities [28]. To further extend this concept, the OWASP CycloneDX framework expands the SBOM model to include Hardware Bill of Materials (HBOM), Operations Bill of Materials (OBOM), and relevant security data such as Vulnerability Disclosure Reports (VDR) and Vulnerability Exploitability Exchange (VEX). Incorporating a broader range of standards, such as the CPE format defined in NIST Interagency Report 7695 [29], would allow software components to be mapped against known vulnerabilities using the CVE system [30]. While the MDS2 2019 version already addresses SBOM, these extensions are not yet integrated.

Future MDS2 versions could provide significant security benefits by fully integrating SIEM/SOAR systems with frameworks such as SIGMA, CPE, and SBOM standards. This would enable healthcare organizations to address vulnerabilities proactively, minimizing potential security risks. The following enhancements could be considered for the next revision cycle of MDS2:

- Integration of SIGMA and SOAR rules for enhanced automated threat detection and response.
- Support for extended BOMs (e.g., HBOM, OBOM) with "End of Life" overviews for software components.
- Adoption of CPE nomenclature for mapping software versions to CVEs, enhancing vulnerability management.
- Implementation of standardized security tests and regular security checks.
- Centralized organization and access to security-related documentation and updates.

Table 2 outlines proposed questions and corresponding category assignments for the MDS2 update.

Table 2: **Overview of recommended new security questions for the next MDS2 revision cycle**

Cyber Threat Detection Rules - CTDR	
CTDR-1	Are there rules for a SIEM System to protect and monitor the device?
CTDR-2	Are there rules to monitor the network traffic with a low false-positive rate (SIGMA Format)?
CTDR-3	Are there rules to instruct a SOAR to automatically protect the medical device?
CTDR-4	Does the device have special warning messages for a SIEM if self-checks fail?
CTDR-5	Does the device have special warning messages for a SIEM if life-threatening settings have been sent to the device?
Cyber Threat Intelligence Information - CTII	
CTII-1	Which CPE notations were determined for the hard- and software components?
CTII-2	Where is the central point to get further security advisories in the future?
CTII-3	What software end-of-life data is already known?
CTII-4	Do you support "minor" and "major" updates to maintain the security process permanently?
Cyber Security Tests and Checks - CSTC	
CSTC-1	Have cyber security tests already been carried out on the device?
CSTC-2	Have standard cyber security checks already been carried out on the device?
CSTC-3	Are there official information about the tests and checks which were published?
Digital Forensic and Incident Response - DFIR	
DFIR-1	Does the device have the ability to provide a triage package for a forensic investigation?
DFIR-2	Does the device have the ability to interact with an Endpoint Detection and Response (EDR) software?
DFIR-3	Are there any complications with active network and endpoint scanning?
Automatic Logoff - ALOF	
ALOF-3	Does the device have a display deactivation function after a certain time so that it does not show any information?
Authorization - AUTH	
AUTH-6	Does the device have a limit to the number of login attempts in the system?
AUTH-7	Does the device have a delay function between failed login attempts in the system?

Regular reviews of the MDS2 framework could help to ensure that it remains current and relevant, providing medical device manufacturers and operators with reliable and actionable security guidance. Although the revision cycle has been reduced from nine to six years, further shortening these intervals is recommended. Introducing both "minor" and "major" updates throughout the year could help address emerging threats more dynamically.

4 Discussion

In this work, we analyzed 147 MDS2 documents from three distinct versions to assess their machine-readability and the feasibility of deriving a holistic landscape of medical device IT security. We performed intra-version comparisons, when applicable, for the 2013 and 2019 versions, as well as inter-version comparisons to evaluate similarities and differences across versions.

From the intra-version comparison, valuable insights emerged. For instance, the finding that just about half of the devices could enforce password policies provides an explanation for the results of [8], which identify weak or default passwords as a major vulnerability of medical devices. Notably, the latest MDS2 version has improved the potential for such evaluations by offering enhanced informative value, which makes future evaluations promising.

Our inter-version analysis revealed substantial inconsistencies in the phrasing of security-related questions between versions, complicating the comparison process. This issue was particularly pronounced when comparing the 2008 version to later versions, as it contained narrowly defined questions that focused primarily on specific types of data, such as ePHI. Although we addressed these issues by employing categorization, the observed deterioration in results between the 2008 and 2013 versions likely reflects this inconsistency rather than a decline in IT security. However, the decline in categories such as remote service connections and remote security updates cannot be attributed to these limitations, as the questions remained largely unchanged between 2008 and 2019. Instead, we suspect this may be due to improved network protection and segmentation on the operator's side, which, while reducing exposure and attack surfaces, also complicates remote security updates and services for manufacturers.

What makes our research approach unique are novel insights that were previously unattainable in landscape analyses. One considerable finding involves the proportion of medical devices that use encrypted connections for data transmission. Until now, no reliable data was available regarding whether the network traffic of medical devices is encrypted. This gap in knowledge hindered the development of attack detection methods in hospital settings, forcing researchers to make assumptions in their studies [31]. Our work demonstrated that, although the use of encrypted connections has steadily increased since the initial MDS2 version, just over half of the devices encrypted network traffic according to the latest MDS2 version documents.

Another critical finding relates to the logging capabilities of medical devices, which are crucial for developing real-time attack detection systems. While the ability to generate log data is essential, it is insufficient unless the device can also forward these logs to a SIEM system. Since the 2019 MDS2 version, the documents explicitly inquire about this capability. According to 2019 data, despite over 80% of devices generating logs, only 11% were capable of forwarding logs to a SIEM instance. This demonstrates the need for further improvements or log data-agnostic approaches to attack detection.

During the analysis, some limitations of MDS2 docu-

ments were encountered. One primary issue is the use of PDF formats, which introduce inaccuracies during data input and impede subsequent machine processing, reducing the documents' evaluability. Additionally, the decentralized distribution of documents leads to inconsistencies and discrepancies between documents and the devices they describe. The MDS2 documents reflect only a snapshot of the device's state at the time of creation. Updates to the device introduced via software patches may not be captured unless the documents are updated and reliably distributed. These factors highlight the need for a structured and standardized method for creating and maintaining MDS2 documents. One potential solution is the establishment of a central archive for MDS2 documents. This archive would ensure the documents are up-to-date and maintain a high level of quality. Manufacturers could update their documents as needed through a version-controlled system, enabling transparency and traceability. The Common Security Advisory Framework (CSAF) could serve as a model for this, as CSAF advisories are created in machine-readable JSON format and can be visualized using dedicated tools [32–34].

Centralized storage would also resolve existing issues, such as inconsistencies in publication dates within document metadata. Accurate publication dates and document versions are crucial for comprehensive landscape analyses. Combined with the ability to scale document use, this approach would improve data quality and enhance the efficiency of document processing [35]. Such an infrastructure aligns with HIMSS' original vision of establishing a centralized infrastructure for MDS2 documents.

Furthermore, the potential exists not only to create an overarching IT security landscape for medical devices but also to enable operators to select specific MDS2 documents for tailored analysis. This would allow them to build security postures according to organizational needs, thereby increasing the relevance and practical utility of the data.

5 Conclusion and Future Work

Operators are already using MDS2 documents to quickly assess the security capabilities of individual medical devices without requiring direct interaction with them. In this study, we propose a novel approach to analyzing the medical device IT security landscape through a systematic assessment of these documents. Given the long lifespan of medical devices, evaluating older documents remains relevant, as many devices are still in operation. The literature indicates life cycles of 13 years for anesthesia machines and ventilators, 14 years for defibrillators, and 16 years for heart-lung machines [36]. Moreover, obtaining a comprehensive understanding of the evolving IT security landscape over time aids in decision-making and risk assessment.

Our findings suggest that intra-version comparisons of MDS2 documents become feasible starting from the second version. Inter-version comparisons, while possible based on defined categories, face limitations, especially when analyzing first-version documents. The security insights derived from MDS2 documents can substantially enhance research and practical efforts in attack detection, particularly in hospital environments with network-enabled medical devices. MDS2 documents address critical questions that were previously challenging to answer, including capabilities for logging, data encryption, and the availability of SBOMs.

It is important to note that this study did not evaluate the accuracy of the information provided in MDS2 documents, i.e., whether the manufacturers' claims accurately reflect the devices' security characteristics. Since these documents are typically delivered to operators at the time of purchase, updates to the devices may not be reflected in the MDS2 documents. The decentralized distribution of these documents as PDFs or Excel files could lead to outdated revisions or updates not being communicated to all users.

Future work should focus on investigating the accuracy of the information in MDS2 documents, aiming to determine the extent to which updates are neglected or revisions are not properly distributed to all stakeholders. Moreover, creating a centralized repository for MDS2 documents, preferably in a structured, machine-readable format, could streamline the distribution of document updates and allow operators to generate individualized IT security status reports for their organizations. This would enhance transparency regarding the IT security status of the medical devices in use, benefiting operators, researchers, and patients.

References

[1] WHO, "Attacks on healthcare worldwide 2023," publication Title: Statista. [Online]. Available: https://www.statista.com/statistics/1303742/number-of-reported-attacks-against-healthcare-worldwide/

[2] B. für Sicherheit in der Informationstechnik, "Die Lage der IT-Sicherheit in Deutschland 2023," Sep. 2021, publication Title: Bundesamt für Sicherheit in der Informationstechnik. [Online]. Available: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Lageberichte/Lagebericht2023.html?nn=129410

[3] "ENISA Threat Landscape 2021," Oct. 2021, publication Title: ENISA Type: Report/Study. [Online]. Available: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2021

[4] "Health Threat Landscape — ENISA," May 2023. [Online]. Available: https://www.enisa.europa.eu/publications/health-threat-landscape

[5] U.S. Department of Health and Human Services, "Healthcare sector cybersecurity," 2023.

[6] A. D. Stern, W. J. Gordon, A. B. Landman, and D. B. Kramer, "Cybersecurity features of digital medical devices: an analysis of fda product summaries," *BMJ open*, vol. 9, no. 6, p. e025374, 2019.

[7] L. Bracciale, P. Loreti, and G. Bianchi, "Cybersecurity vulnerability analysis of medical devices purchased by national health services," *Scientific Reports*, vol. 13, no. 1, p. 19509, 2023.

[8] C. M. Mejía-Granda, J. Fernández-Alemán, J. M. Carrillo de Gea, and J. García Berná, "Security vulnerabilities in healthcare: an analysis of medical devices and software," *Medical & Biological Engineering & Computing*, vol. 62, Oct. 2023.

[9] allianz-fuer cybersicherheit, "Sicherheit von Medizinprodukten - Leitfaden zur Nutzung des MDS2 aus 2019," Nov. 2019.

[10] Z. Wang, P. Ma, X. Zou, J. Zhang, and T. Yang, "Security of medical cyber-physical systems: An empirical study on imaging devices," in *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2020, pp. 997–1002.

[11] A. Alhammad, M. M. Yusof, and D. I. Jambari, "A review of cyber threats to medical devices integration with electronic medical records," in *2022 International Conference on Cyber Resilience (ICCR)*, 2022, pp. 1–6.

[12] T. Yaqoob, H. Abbas, and M. Atiquzzaman, "Security Vulnerabilities, Attacks, Countermeasures, and Regulations of Networked Medical Devices—A Review," *IEEE Communications Surveys Tutorials*, vol. 21, no. 4, pp. 3723–3768, 2019.

[13] "IEC/TR 80001-2-2:2012," publication Title: ISO. [Online]. Available: https://www.iso.org/standard/57939.html

[14] J. T. Force, "Security and Privacy Controls for Information Systems and Organizations," National Institute of Standards and Technology, Tech. Rep. NIST Special Publication (SP) 800-53 Rev. 5, Dec. 2020. [Online]. Available: https://csrc.nist.gov/pubs/sp/800/53/r5/upd1/final

[15] I. O. for Standardization, *Information Security, Cybersecurity and Privacy Protection: Information Security Controls*. International Organization for Standardization, 2022.

[16] CISA. Require strong passwords | CISA. [Online]. Available: https://www.cisa.gov/secure-our-world/require-strong-passwords

[17] Orange. Cyber attacks in healthcare sector worldwide by type 2022. [Online]. Available: https://www.statista.com/statistics/1362863/cyber-attacks-on-healthcare-organizations-worldwide-by-type/

[18] HIPAA Journal. Causes of u.s. healthcare data breaches in 2022. [Online]. Available: https://www.statista.com/statistics/1274643/causes-of-us-healthcare-data-breaches/

[19] Brandt, Mathias. Infographic: Security risk software. [Online]. Available: https://www.statista.com/chart/1859/companies-by-the-number-of-security-vulnerabilities

[20] NIST, "NVD - CPE." [Online]. Available: https://nvd.nist.gov/products/cpe

[21] SecurityScorecard. Vulnerable software by product: Windows. [Online]. Available: https://www.cvedetails.com/product-search.php?vendor_id=0&search=linux&operatingsystem=1

[22] L. Johnson, K. Dempsey, R. Ross, S. Gupta, and D. Bailey, "Guide for security-focused configuration management of information systems," National Institute of Standards and Technology, Tech. Rep. NIST Special Publication (SP) 800-128, 2019. [Online]. Available: https://csrc.nist.gov/pubs/sp/800/128/upd1/final

[23] ISO Central Secretariat. ISO/IEC 27001:2022. [Online]. Available: https://www.iso.org/standard/27001

[24] FDA, "Cybersecurity in medical devices: Quality system considerations and content of premarket submissions." [Online]. Available: https://www.fda.gov/media/119933/download

[25] K. Scarfone and M. Souppaya, "Cybersecurity log management planning guide," National Institute of Standards and Technology, Tech. Rep. NIST SP 800-92r1 ipd, 2023. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-92r1.ipd.pdf

[26] J. Boyens, A. Smith, N. Bartol, K. Winkler, A. Holbrook, and M. Fallon, "Cybersecurity supply chain risk management practices for systems and organizations," National Institute of Standards and Technology, Tech. Rep. NIST Special Publication (SP) 800-161 Rev. 1, 2022. [Online]. Available: https://csrc.nist.gov/pubs/sp/800/161/r1/final

[27] "Improving the Nation's Cybersecurity," May 2021, publication Title: Federal Register. [Online]. Available: https://www.federalregister.gov/documents/2021/05/17/2021-10460/improving-the-nations-cybersecurity

[28] Medical Device Cybersecurity Working Group, "Principles and practices for software bill of materials for medical device cybersecurity." [Online]. Available: https://www.imdrf.org/sites/default/files/2023-04/Principles%20and%20Practices%20for%20Software%20Bill%20of%20Materials%20for%20Medical%20Device%20Cybersecurity%20%28N73%29.pdf

[29] B. A. Cheikes, D. Waltermire, and K. Scarfone, "Common platform enumeration : naming specification version 2.3," National Institute of Standards and Technology, Gaithersburg, MD, Tech. Rep. NIST IR 7695, 2011. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7695.pdf

[30] I. T. L. Computer Security Division. Security content automation protocol | CSRC | CSRC. [Online]. Available: https://csrc.nist.gov/projects/security-content-automation-protocol

[31] S. B. Weber, S. Stein, M. Pilgermann, and T. Schrader, "Attack detection for medical cyber-physical systems–a systematic literature review," *IEEE Access*, vol. 11, pp. 41 796–41 815, 2023, conference Name: IEEE Access. [Online]. Available: https://ieeexplore.ieee.org/document/10107991

[32] "Common Security Advisory Framework (CSAF)," publication Title: Bundesamt für Sicherheit in der Informationstechnik. [Online]. Available: https://www.bsi.bund.de/DE/Themen/Unternehmen-und-Organisationen/Informationen-und-Empfehlungen/Empfehlungen-nach-Angriffszielen/Industrielle-Steuerungs-und-Automatisierungssysteme/CSAF/CSAF.html?nn=954494

[33] "CSAF 2.0 JSON Schema Viewer." [Online]. Available: https://json.csaf.io/

[34] "Secvisogram CSAF 2.0 Editor." [Online]. Available: https://secvisogram.github.io/

[35] Yumpu.com, "Manufacturers Disclosure Statement for Medical Device ... - himss," publication Title: yumpu.com. [Online]. Available: https://www.yumpu.com/en/document/read/22293145/manufacturers-disclosure-statement-for-medical-device-himss

[36] G. Seo, S. Park, and M. Lee, "How to calculate the life cycle of high-risk medical devices for patient safety," *Front Public Health*, vol. 10, p. 989320, 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9515981/

Chapter 7

SzA4Hosp – Attack Detection in Hospitals

This chapter gives an overview of the contributions and the impact of the white paper Weber et al. (2024a):

Simon Weber, Michael Pilgermann, Stefan Stein, Thomas Schrader

"SzA4Hosp — Systeme zur Angriffserkennung in der Medizinischen Versorgung"

White paper made available via the publication server of the Brandenburg University of Applied Sciences[1] and submitted to the BAK MV.

7.1 Summary

This white paper addresses the implementation of SzA within German hospitals. Induced by the IT Security Act of May 2021, which requires hospitals to adopt SzA, it provides insights into the present state of SzA adoption, regulatory compliance, and the challenges hospitals face when integrating these systems.

Data was gathered through a nationwide survey of hospital operators, supplemented by expert interviews and collaborative workshops, to evaluate the present state of SzA implementation and to discuss industry-specific challenges. The survey revealed substantial variation in SzA readiness, with many hospitals reporting limited progress, especially in the areas concerning MT and VT. The workshops and interviews offered additional insights into these findings, highlighting obstacles such as limited financial and technical resources, complexity, and the industry's dependency on a standardized approach. The report also examines established best practices from national and international frameworks, to provide practical recommendations for enhancing SzA in hospitals. Recommendations include the phased integration of monitoring and logging practices and the development of clear incident response protocols to improve detection capabilities. Presenting these strategies, the report aims to equip the industry with specific and actionable recommendations.

[1] https://doi.org/10.25933/opus4-3207

7.2 Personal Contribution

Simon Weber developed the survey questionnaire, which he and Michael Pilgermann subsequently revised together. Simon Weber conducted the survey evaluation. He also designed the framework for the semi-structured interviews and conducted them together with Michael Pilgermann. In addition, the authors took part in various BAK meetings and workshops in order to present interim results and obtain direct feedback from hospital representatives. The review of relevant standards and best practices was undertaken jointly by Simon Weber and Michael Pilgermann, who assessed their applicability to the industry-specific challenges. Together, they derived the findings and formulated actionable recommendations. Stefan Stein and Thomas Schrader contributed feedback and insights during discussions. Both reviewed and provided input on drafts of the manuscript, which were co-authored by Simon Weber and Michael Pilgermann.

7.3 Importance and Impact on this Thesis

This white paper serves as a foundational component of this thesis by providing a comprehensive assessment of the present state of attack detection in hospitals, which has been essential in understanding the industry's unique challenges. It includes an in-depth analysis of national and international standards and best practices, along with an evaluation of the extent to which general recommendations are applicable to hospitals, identifying areas where specific solutions are required. Findings from the entire period of this doctorate have been directly incorporated into this white paper, resulting in the formulation of 44 recommendations for attack detection in hospitals. These recommendations form the basis for updating the B3S for hospitals, demonstrating the practical applicability and relevance of this research to Germany's hospitals. The report substantiates the immediate impact and added value of this dissertation for enhancing resilience in the German healthcare sector. Intended as a direct working basis for the BAK MV, this study was written in German and integrated as such into this dissertation.

SzA4Hosp – Systeme zur Angriffserkennung in der Medizinischen Versorgung

SIMON B. WEBER
Heinrich-Heine-University
Düsseldorf, Germany
Simon.Weber@hhu.de

MICHAEL PILGERMANN
University of Applied Sciences
Brandenburg a. d. Havel, Germany
Michael.Pilgermann@th-brandenburg.de

STEFAN STEIN
University of Applied Sciences
Brandenburg a. d. Havel, Germany
Stefan.Stein@th-brandenburg.de

THOMAS SCHRADER
University of Applied Sciences
Brandenburg a. d. Havel, Germany
Thomas.Schrader@th-brandenburg.de

Abstract

Dieser Abschlussbericht stellt das Ergebnis der Projektarbeit über die Implementierung von Systemen zur Angriffserkennung (SzA) in deutschen Krankenhäusern im Kontext des IT-Sicherheitsgesetzes 2.0 und des branchenspezifischen Sicherheitsstandards (B3S) für die medizinische Versorgung dar. Ziel des Projekts war es, den aktuellen Umsetzungsstand von SzA in deutschen Krankenhäusern zu analysieren und Handlungsempfehlungen für die Weiterentwicklung des B3S zu erarbeiten. Die Analyse basiert auf einer umfangreichen Befragung von Krankenhausbetreibern, Expertengesprächen sowie der Auswertung relevanter nationaler und internationaler Standards und Good Practices. Die Ergebnisse zeigen deutliche Unterschiede im Reifegrad der SzA-Implementierung zwischen verschiedenen Bereichen, wobei die Informationstechnik branchenweit am fortgeschrittensten ist. Der Bericht bietet konkrete Vorschläge zur Verbesserung der IT-Sicherheitslage in Krankenhäusern und betont die Notwendigkeit kontinuierlicher Weiterentwicklungen der SzA-Systeme, um den steigenden Anforderungen der IT-Sicherheit in der stationären Versorgung gerecht zu werden.

1 Einleitung

Am 18. Mai 2021 wurde das „Zweite Gesetz zur Erhöhung der Sicherheit informationstechnischer Systeme" [1] (IT-Sicherheitsgesetz 2.0) veröffentlicht und trat somit am Folgetag in Kraft. Es knüpft an die Pflichten für Betreiber Kritischer Infrastrukturen aus dem „Gesetz zur Erhöhung der Sicherheit informationstechnischer Systeme" [2] (IT-Sicherheitsgesetz, 2015) an und konkretisiert die Auflage zur Einhaltung „angemessener organisatorischer und technischer Vorkehrungen" dahingehend, dass diese ab dem 1. Mai 2023 auch den Einsatz von Systemen zur Angriffserkennung (SzA) umfassen muss.
Das Bundesamt für Sicherheit in der Informationstechnik (BSI) hat am 26.09.2022 die „Orientierungshilfe zum Einsatz von Systemen zur Angriffserkennung" [3] (OH SzA) veröffentlicht und den Betreibern ein Hilfsmittel an die Hand gegeben, um die neue Anforderung zu SzA sachgerecht umzusetzen. Dem BSI kommt hierbei eine besondere Rolle zu, da die Betreiberorganisationen gegenüber dem BSI regelmäßig nachweisen müssen, dass beim Betrieb ihrer Kritischen Infrastrukturen angemessene organisatorische und technische Vorkehrungen – ab Mai 2023 inklusive jener zu SzA – umgesetzt sind.
Bereits im IT-Sicherheitsgesetz von 2015 [2] wurde den Betreiberorganisationen und ihren Branchenverbänden die Möglichkeit eröffnet, branchenspezifische Sicherheitsstandards (B3S) vorzuschlagen, deren Eignung zur Umsetzung der angemessenen organisatorischen und technischen Vorkehrungen vom BSI sodann auf Antrag geprüft und festgestellt wird. Damit sollen die Betreiber Handlungssicherheit hinsichtlich der Einhaltung des IT-Sicherheitsgesetzes (§8a) erhalten, sofern sie einen für sie einschlägigen B3S umgesetzt haben. Für die Krankenhäuser wurde von dieser Möglichkeit zur Einreichung von B3S Gebrauch gemacht – das BSI hat zuletzt den „Branchenspezifischen Sicherheitsstandard „Medizinische Versorgung"" in der Version 1.2 vom 08.12.2022 [4] als geeignet festgestellt.
Die Deutsche Krankenhausgesellschaft e.V. (DKG) hat in dieser regulatorischen und normativen Gemengelage das Projektteam von der Technischen Hochschule Brandenburg (THB) beauftragt, um bei der Fortschreibung des B3S hinsichtlich der Anteile zur Umsetzung der Anforderung an SzA zu unterstützen. Dieses Dokument stellt den Abschlussbericht zu dieser Unterstützungsleistung dar.

1.1 Zielsetzung

Zwischen dem Auftraggeber DKG (AG) und dem Auftragnehmer THB (AN) wurden für das Projekt die folgenden Ziele vereinbart:

- Begleitung der Fortschreibung des „B3S für die Gesundheitsversorgung im Krankenhaus" auf die neue Hauptversion 2.0
- Konkretisierung der Anforderungen nach BSIG und OH SzA für die Branche bezüglich der Systeme zur Angriffserkennung

Die Zielsetzung des Projektes beschränkt sich auf die technischen und prozessualen Gegebenheiten beim Betrieb von Krankenhäusern hinsichtlich der Systeme zur Angriffserkennung. Eine Untersuchung von Ursachen für aktuelle Umsetzungsstände und somit die Einbettung der Thematik in die übergreifende

wirtschaftliche Situation von Krankenhäusern beispielsweise im Zusammenhang mit der Krankenhausreform oder den Besonderheiten bei der Finanzierung von Investitionen, bei welcher Entscheidungen in Abhängigkeit von der Trägerform nicht immer abschließend vom Betreiber getroffen werden, ist explizit nicht Bestandteil des Projektes.

1.2 Stakeholder

Die nachfolgenden Organisationen und Gruppen sind für die Durchführung dieses Projektes und insofern im weiteren Sinne für die Ausgestaltung von SzA für Krankenhäuser in Deutschland besonders relevant:

Die **Deutsche Krankenhausgesellschaft e.V. (DKG)** fungiert als Auftraggeber für dieses Projekt. Laut eigener Aussage[1] steht die Deutsche Krankenhausgesellschaft (DKG) für 28 Mitgliedsverbände von Krankenhausträgern. Die DKG ist in mehrfacher Hinsicht in das Thema SzA und B3S aktiv eingebettet: Die DKG ist Bereitsteller des einzigen Branchenspezifischen Sicherheitsstandards für den Betrieb von Krankenhäusern in Deutschland und stimmt sich im Rahmen der Erstellung selbiger sowohl mit dem Branchenarbeitskreis Medizinische Versorgung sowie den zuständigen Gremien der Deutschen Krankenhausgesellschaft ab.

Die **Technische Hochschule Brandenburg (THB)** ist mit ihrem Forschungsschwerpunkt „Interdisziplinäre Sicherheitsforschung" auf der Forschungslandkarte der Hochschulrektorenkonferenz[2] offiziell gelistet. Im Fachbereich Informatik und Medien der THB gibt es ausgeprägte Forschungs- und Lehrtätigkeiten sowohl in der IT-Sicherheit als auch in der Medizininformatik. Die Schnittmengen der beiden Themengebiete werden seit 2021 von Forschenden des Fachbereichs in der MedSec-Gruppe[3] aktiv bearbeitet. Mitwirkende in der MedSec-Gruppe fungieren als Auftragnehmer dieses Projektes.

Der **UP KRITIS**[4] ist seit 2007 eine Kooperation zwischen Betreibern Kritischer Infrastrukturen und relevanten Behörden in Deutschland. Die Gremienstruktur innerhalb des UP KRITIS sieht vor, dass für branchenspezifische Themen und Ausgestaltungen sogenannte Branchenarbeitskreise (BAK)[5] gebildet werden. Für die Krankenhäuser in Deutschland stellt der Branchenarbeitskreis „Medizinische Versorgung" diese entsprechende Ausrichtung zur Verfügung. Zudem werden branchenübergreifende Spezialthemen in sogenannten Themenarbeitskreisen (TAK)[6] bearbeitet. Für das Thema „Systeme zur Angriffserkennung" ist insbesondere der TAK „Detektion" relevant.

---

[1] https://www.dkgev.de/dkg/aufgaben-ziele/
[2] https://www.hrk.de/themen/forschung/forschungslandkarte/
[3] https://informatik.th-brandenburg.de/forschung-und-kooperation/schwerpunkte-und-themen/medsec-it-sicherheit-in-der-medinischen-versorgung/
[4] https://www.bsi.bund.de/DE/Themen/KRITIS-und-regulierte-Unternehmen/Kritische-Infrastrukturen/UP-KRITIS/up-kritis_node.html
[5] https://www.bsi.bund.de/DE/Themen/KRITIS-und-regulierte-Unternehmen/Kritische-Infrastrukturen/UP-KRITIS/Branchenarbeitskreise/branchenarbeitskreise_node.html
[6] https://www.bsi.bund.de/DE/Themen/KRITIS-und-regulierte-Unternehmen/Kritische-Infrastrukturen/UP-KRITIS/Themenarbeitskreise/themenarbeitskreise_node.html

Das **Bundesamt für Sicherheit in der Informationstechnik (BSI)** fungiert in erster Linie als Aufsichtsbehörde über Betreiber Kritischer Infrastrukturen in Deutschland bezüglich der Umsetzung der Regelungen zu IT-Sicherheit nach dem BSI-Gesetz. So prüft es beispielsweise die von den Branchen eingereichten B3S auf Eignung zur Einhaltung der IT-Sicherheitsauflagen und überwacht ebenfalls die Nachweise zur Umsetzung der Regeln, welche von den Betreibern alle zwei Jahre überlassen werden müssen. Das BSI gibt ebenfalls konkretisierende Handreichungen heraus, wie das BSI-Gesetz umgesetzt werden soll. Dies erfolgt regelmäßig im Format von „Orientierungshilfen". Letztendlich betreibt das BSI auch die Geschäftsstelle des UP KRITIS.

1.3 Ausgestaltung und Vorgehen

Das Projekt wurde im Zeitraum Mai 2023 bis Juni 2024 durchgeführt. Für die Durchführung des Projektes wurden unterschiedliche Methoden herangezogen:

- Unterlagen-Analyse: Heranziehung von für den Stand der Technik relevanten Gesetzen, untergesetzlichen Regelungen, Standards und Good Practices.
- Fragebogen: Erhebung von Umsetzungsständen zu Systemen zur Angriffserkennung bei Betreibern von Krankenhäusern in Deutschland.
- Workshops: Validierung von Thesen aus der Erhebung zur Umsetzung sowie Einordnung und Auflösung von statistischen Abweichungen.
- Experten-Interviews: Semistrukturierte Interviews mit ausgewählten Ansprechpartnern für konkrete Rückkopplungen aus der Setzung von Standards (sowohl Orientierungshilfen als auch B3S) sowie aus der konkreten Umsetzung selbiger bei den Betreibern.
- Teilnahme an Arbeitsgruppentreffen: Präsentation von Zwischenständen und Einholung von Feedback für die Finalisierung.

Zudem wurden Erkenntnisse der MedSec-Gruppe aus weiterführenden F&E-Aktivitäten, insbesondere hinsichtlich der Analyse von MDS2-Dokumenten, sowie Erfahrungswerte aus Aufbau von und Mitwirkung in IT-Sicherheitsteams mit herangezogen und sind so mittelbar in das Projektergebnis mit eingeflossen.

Wie in Abbildung 1 dargestellt, erfolgte die Bearbeitung in drei Arbeitspaketen:

- Bestandsaufnahme,
- Ableitung von Handlungsempfehlungen und
- Ableitung von Formulierungsvorschlägen.

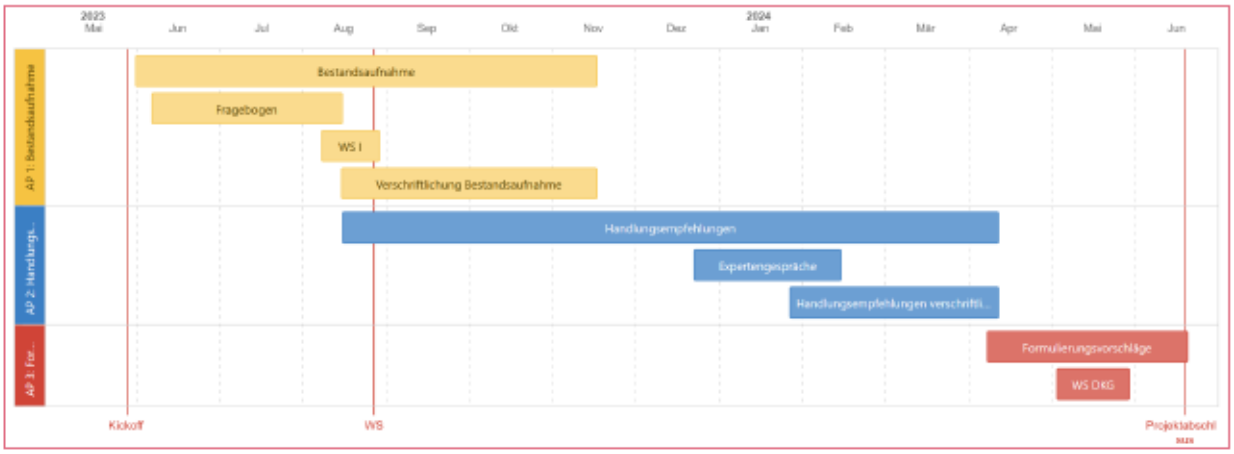

Abbildung 1: Projektplan mit drei Arbeitspaketen

Die Tätigkeiten in Arbeitspaket 2 (Handlungsempfehlungen), insbesondere hinsichtlich der Analyse einschlägiger nationaler und internationaler Dokumente, bedurften deutlich ausführlicherer Analyse als ursprünglich geplant. Der Zeitrahmen von Arbeitspaket 2 und Arbeitspaket 3 wurde daher gegenüber der ursprünglichen Planung um ca. 12 Wochen verlängert.

2 Bestandsaufnahme

Die Bestandsaufnahme zu SzA beim Betrieb von Krankenhäusern in Deutschland ist in den nachfolgenden Unterabschnitten dokumentiert:

- Eine initiale Kurzeinführung gibt einen Überblick über die wichtigsten Technologien zur Angriffserkennung.
- Im Anschluss ist ein Einblick in Zahlengerüste zu den Betreibern von Krankenhäusern in Deutschland dokumentiert
- Der nachfolgende Unterabschnitt beleuchtet den konkreten Regelungsstand in Deutschland beim Betrieb Kritischer Infrastrukturen und gibt einen kurzen Ausblick auf aktuelle diesbezügliche Entwicklungen.
- In Unterabschnitt 2.4 wird sodann auf den aktuellen Stand zur Erstellung der Branchenspezifischen Sicherheitsstandards (B3S) eingegangen und die Abbildung von SzA im aktuell gültigen B3S für Krankenhäuser nachgehalten.
- Abschließend ist in Unterabschnitt 2.5 eine Klärung der Begrifflichkeit „Stand der Technik" dokumentiert.

Im Sommer 2023 wurde vom Projektteam eine Umfrage unter deutschen Krankenhäusern zum Thema SzA durchgeführt und ausgewertet. Die Durchführung der Umfrage und Ergebnisse aus der selbigen sind in einem eigenen Abschnitt (3 Betreiberbefragung Umsetzung SzA) zusammengetragen.

2.1 Technologien zur Angriffserkennung

Erste Technologien zur Angriffserkennung gab es bereits in den 1980ern. Abbildung 2 ordnet die wichtigsten technologischen Entwicklungen zeitlich ein.

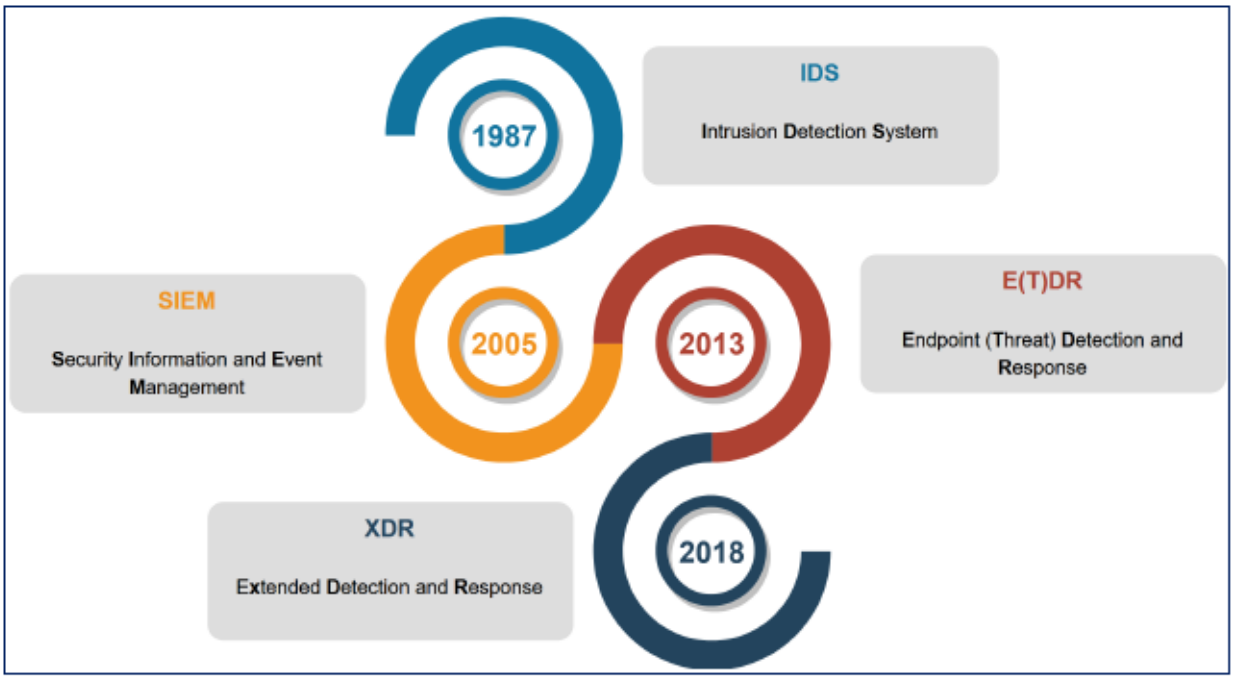

Abbildung 2: Historische Entstehung von Angriffserkennungssystemen (Quelle: [5])

Nachfolgend werden diese kurz beschrieben, da ein diesbezügliches Grundverständnis für die weitere Befassung mit dem Thema zuträglich ist:

- Intrusion Detection System (Einbruchserkennungssystem, IDS): Intrusion Detection Systeme basierten ursprünglich nur auf Signaturen, um Vorfälle zu erkennen. Später wurden anomaliebasierte IDS entwickelt, die Abweichungen von einem zuvor definierten Normalverhalten erkennen. Grundsätzlich lassen sich IDS in zwei Kategorien unterteilen: Je nachdem, ob ein IDS den Netzverkehr oder einzelne Geräte (Hosts) überwacht, werden sie als NIDS (Network Intrusion Detection Systems) oder HIDS (Host Intrusion Detection Systems) kategorisiert.
- Security Information and Event Management (SIEM): Diese Systeme sammeln und kategorisieren Daten (vorzugsweise Protokollierungsdaten) aus möglichst vielen Quellen innerhalb eines Netzwerks und korrelieren sie anschließend, um Angriffe zu erkennen.
- Endpoint Detection and Response (EDR): EDR-Systeme schützen Endpunkte (PCs, IoT, Mobiltelefone, etc.) eines Netzwerks. Dazu werden oft sogenannte Agenten (Programme, die auf den Endgeräten laufen und Informationen zu Verhalten und Ereignissen an eine zentrale Stelle weiterleiten) eingesetzt, welche bei Auffälligkeiten Alarm schlagen.
- Extended Detection and Response (XDR): XDR verbindet die Ansätze von SIEM und EDR zu einem und sammelt so viele relevante Informationen wie möglich aus der gesamten IT-Infrastruktur (Endpoints, Server, EDR, Cloud-Anwendungen, Netzwerkkomponenten, etc.), um Angriffe zu erkennen. [5]

2.2 Betreiberübersicht Krankenhäuser in Deutschland

Der Zwischenbericht aus dem Digitalradar von September 2022 bietet eine für die Zwecke dieses Projektes präzise Betrachtung der Krankenhauslandschaft Deutschlands. [6] Unter Zusammenführung von Bericht und Grunddaten zu Krankenhäusern in Deutschland des Statistischen Bundesamtes aus dem Jahr 2020 [7] wurden in 2020 1903 Krankenhäuser betrieben, welche wiederum 487 783 Betten für die stationäre Versorgung zur Verfügung stellten. Bei der Trägerschaft wird zwischen den folgenden drei Varianten unterschieden:

- Öffentlicher Träger (29 % der Krankenhäuser)
- Freigemeinnütziger Träger (33 % der Krankenhäuser)
- Privater Träger (38 % der Krankenhäuser)

Vergleicht man die Aufteilung von Krankenhäusern mit der Anzahl von Krankenhausbetten, ergibt sich ein anderes Bild; Grund hierfür ist die Tatsache, dass öffentliche Krankenhäuser im Durchschnitt eine dreimal größere Kapazität als private Krankenhäuser bereitstellen.

Der Zwischenbericht hat im Rahmen der Auswertung von EMRAM-Reifegraden (Electronic Medical Record Adoption Model) auch die Dimension der „IT- & Informationssicherheit" (INFO) beleuchtet. Die

durchschnittliche Spannbreite bei den Erfüllungsgraden dieser Dimension liegt bei 22 % bis 100 % und ist demnach höher als bei den anderen EMRAM-Dimensionen.

2.3 Regulatorischer und normativer Status quo zu IT-Sicherheit in Krankenhäusern

Im Juli 2015 wurde das „Gesetz zur Erhöhung der Sicherheit informationstechnischer Systeme" [2] (IT-Sicherheitsgesetz) in Kraft gesetzt. Damit wurde in Deutschland erstmalig auf nationaler Ebene legaldefiniert, was Kritische Infrastrukturen sind. Eine der maßgeblichen Auflagen an Betreiber Kritischer Infrastrukturen wurde in das „Gesetz über das Bundesamt für Sicherheit in der Informationstechnik" (BSIG) als §8a aufgenommen, wonach beim Betrieb dieser Anlagen der Stand der Technik bezüglich der IT-Sicherheit eingehalten werden muss. Mit dem Gesetz wurden ursprünglich sieben Sektoren der Kritischen Infrastrukturen den Regelungen unterworfen, wobei einer dieser Sektoren der Sektor „Gesundheit" ist.

Die konkrete Definition, wann eine Anlage als Kritische Infrastruktur eingestuft ist, wurde per §10 BSIG an eine Rechtsverordnung des Bundesministeriums des Innern delegiert. Diese „Verordnung zur Bestimmung Kritischer Infrastrukturen nach dem BSI-Gesetz" [8] (BSI-Kritisverordnung – BSI-KritisV) wurde 2016 in Kraft gesetzt und einige Male weiterentwickelt. Der Sektor „Gesundheit" ist in §6 dieser Verordnung geregelt – unter den vier einschlägigen Dienstleistungen für diesen Sektor findet sich jene für Krankenhäuser mit folgender Bezeichnung wieder: „die stationäre medizinische Versorgung". Die quantitative Beurteilung, ob eine Anlage zu den Kritischen Infrastrukturen gezählt wird, ist Anhang Teil 3 jener Verordnung zu entnehmen: Demnach wurde in der Dienstleistung „Stationäre medizinische Versorgung" die Anlage „Krankenhaus" mit einem Schwellenwert von 30.000 vollstationären Fällen pro Jahr aufgenommen.

Das BSI hat daraufhin in seinen jährlichen Lageberichten Aussagen zu den betroffenen Anlagen und Betreibern dokumentiert. Da diese Berichterstattung ab 2020 in den Berichten nicht mehr zu finden ist, datierten die letzten diesbezüglichen Angaben auf den Bericht aus 2019. Demnach waren im Jahr 2019 358 Betreiber nachweispflichtig; ca. 1700 KRITIS-Anlagen waren registriert. [9] Seit Mai 2024 veröffentlicht das BSI Statistiken und Kennzahlen zu Kritischen Infrastrukturen auf einer eigens dafür eingerichteten Webseite „KRITIS in Zahlen"[7]. Demnach betrieben zum Stichtag 29.04.2024 1119 Betreiber (davon 210 Sektor Gesundheit) insgesamt 2019 Anlagen der Kritischen Infrastrukturen (davon 330 Sektor Gesundheit).

Im Mai 2021 wurde sodann das „Zweite Gesetz zur Erhöhung der Sicherheit informationstechnischer Systeme" [1] verkündet und in Kraft gesetzt. Für Betreiber Kritischer Infrastrukturen finden sich vergleichsweise wenige Regelungsbestandteile in dem Gesetz wieder. Bezüglich der Verpflichtung zum Treffen angemessener organisatorischer und technischer Vorkehrungen wurde jedoch im BSI-Gesetz ein

[7] https://www.bsi.bund.de/DE/Themen/KRITIS-und-regulierte-Unternehmen/Kritische-Infrastrukturen/KRITIS-in-Zahlen/kritis-in-zahlen.html

neuer Absatz 1a in §8a eingefügt, nach welchem eben diese Vorkehrungen den Einsatz von Systemen zur Angriffserkennung umfassen müssen. Für die Umsetzung dieser Verpflichtung wurde der 1. Mai 2023 als Frist gesetzt.

2.3.1 Stand der Umsetzung IT-Sicherheitsgesetz

Das BSI hatte im ersten Quartal 2023 eine Untersuchung zur Wirksamkeit der IT-Sicherheitsgesetze unter Betreibern Kritischer Infrastrukturen durchführen lassen. Der Ergebnisbericht von April 2023 [10] fasst die Antworten von den 379 KRITIS-Unternehmen, die sich beteiligt hatten, zusammen und gibt nachfolgend Einblick in den aktuellen Umsetzungsstand.

2.3.1.1 Sektorenübergreifender Stand

Übergreifend über alle Sektoren der Kritischen Infrastrukturen fasst der Bericht zusammen, dass die Umsetzung der technischen Sicherheitsmaßnahmen weit vorangeschritten ist, bei der Umsetzung der organisatorischen Maßnahmen (zu denen auch „Security Operations" gezählt wird) die Betreiberorganisationen etwas weiter zurückliegen. Die Einführung von „Security Operations" wird als eines der beiden größten Defizite benannt – nur 60 % der Betreiberorganisationen führen dies durch. Weitere 21 % befinden sich jedoch aktuell in der Planungsphase für Security Operations.

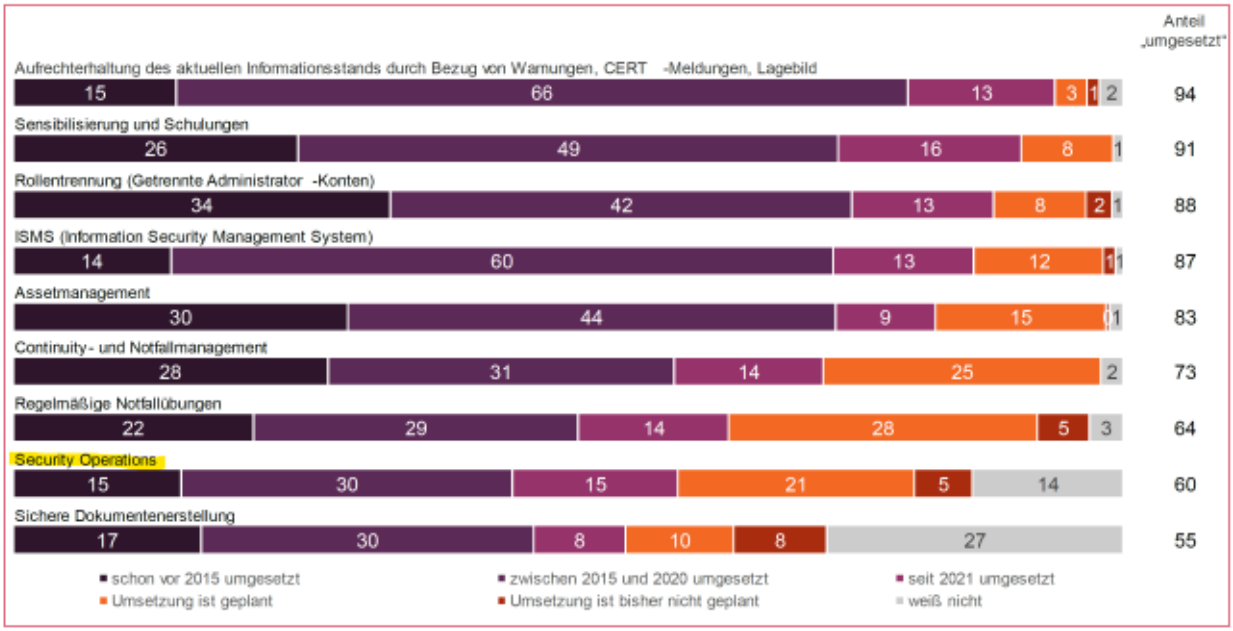

Abbildung 3: Umgesetzte organisatorische Sicherheitsmaßnahmen (Quelle [10])

Gemeinhin wird eine erhöhte IT-Bedrohungslage wahrgenommen; als häufigste Form von Cyberattacken wurden Phishing und Schadsoftware in Mailanhängen benannt. 40 % der teilnehmenden Organisationen mussten in den vergangenen zwei Jahren auf Cyber-Attacken reagieren. „Im Durchschnitt lag der finanzielle Gesamtschaden durch Cyber-Attacken in den letzten zwei Jahren bei 128.000 Euro"

„Als Hauptgründe für die noch unvollständige Umsetzung der gesetzlichen Anforderungen gelten vor allem Personalmangel und fehlende finanzielle Mittel." Dabei seien die Kosten für die Anpassung von IT-Systemen und -prozessen die größte Herausforderung. Zudem arbeitet der Ergebnisbericht einen starken

Zusammenhang zwischen „wirtschaftlicher Leistungsfähigkeit und Umsetzung der gesetzlichen Maßnahmen zur IT-Sicherheit" heraus.
Die Produkte und Publikationen des BSI sind weithin bekannt; bezüglich deren Weiterentwicklung zählen zielgruppen- / branchenspezifische Informationen und Umsetzungshinweise zu den häufigsten Wünschen.
Sofern es für einen Sektor einen B3S gibt, wenden die „weitaus meisten" Unternehmen (80 %) diesen auch an; über zwei Drittel sprechen ihm einen Mehrwert zu.
Laut Ergebnisbericht kann die „Wirksamkeit der IT-Sicherheitsgesetze ... insgesamt als gut bewertet werden". [10]

2.3.1.2 Sektorspezifischer Stand

25 % der 379 antwortenden KRITIS-Unternehmen gehörten dem Sektor Gesundheit an. Eine Aufschlüsselung auf Strukturen innerhalb des Gesundheitssektors (Branchen oder Kritische Dienstleistungen) wurde in dem Bericht nicht dokumentiert.
Dem Gesundheitssektor werden übergreifend Defizite bei der Umsetzung attestiert – von denjenigen Unternehmen, die nur 70 % der Maßnahmen umgesetzt haben, entfallen 45 % auf den Gesundheitssektor.

2.3.2 Besonderheit SGB V

Im Oktober 2020 wurde mit der Inkraftsetzung des „Gesetzes zum Schutz elektronischer Patientendaten in der Telematikinfrastruktur" [11] (Patientendaten-Schutz-Gesetzes, PDSG) auch eine Änderung bei den Regelungen zur IT-Sicherheit beim Betrieb von Krankenhäusern eingeführt: Mittels Änderungsbefehl auf Sozialgesetzbuch V wurde ein neuer § 75c eingeführt, welcher Krankenhäuser verpflichtet, „nach dem Stand der Technik angemessene organisatorische und technische Vorkehrungen zur Vermeidung von Störungen der Verfügbarkeit, Integrität und Vertraulichkeit sowie der weiteren Sicherheitsziele" zu treffen. Für die Umsetzung dieser Verpflichtung wird explizit auf das Mittel der Branchenspezifischen Sicherheitsstandards (B3S) verwiesen, welche hierfür herangezogen werden können.
Am 26.03.2024 ist das Gesetz zur Beschleunigung der Digitalisierung des Gesundheitswesens [12] (Digital-Gesetz, DigiG) in Kraft getreten. Die IT-Sicherheit in Krankenhäusern ist nunmehr in § 391 SGB V geregelt. Faktisch wurde damit § 75 c abgelöst. Absatz 1 des Paragraphen zieht für die IT-Sicherheit wieder den „Stand der Technik" heran. In Absatz 4 wird ausgeführt, dass mit Anwendung eines vom BSI als geeignet festgestellten B3S die Verpflichtungen u.a. nach Stand der Technik insbesondere erfüllt werden. Eine explizite Anforderung zur Umsetzung von Systemen zur Angriffserkennung, wie sie im BSI-Gesetz für Betreiber Kritischer Infrastrukturen existiert, existiert im Digital-Gesetz zumindest für Krankenhäuser nicht.

2.3.3 Ausblick: NIS-Umsetzungsgesetz und KRITIS-Dachgesetz

In Juli und Dezember 2023 hatte das Bundesministerium des Innern Referentenentwürfe eines „Gesetzes zur Umsetzung der CER-Richtlinie und zur Stärkung der Resilienz kritischer Anlagen" [13], [14] (KRITIS-Dachgesetz – KRITIS-DachG) veröffentlicht. Mit diesem Gesetz sollen EU-Vorgaben umgesetzt und offenbar Regelungen zum Schutz Kritischer Infrastrukturen konsolidiert werden. Es konzentriert sich jedoch auf den physischen Schutz und wirkt sich nicht auf das Thema „Systeme zur Angriffserkennung" beim Betrieb von Krankenhäusern aus. Insofern wird es im weiteren Verlauf nicht berücksichtigt.

Im Mai 2024 hat das Bundesministerium des Innern den Referentenentwurf „Entwurf eines Gesetzes zur Umsetzung der NIS-2-Richtlinie und zur Regelung wesentlicher Grundzüge des Informationssicherheitsmanagements in der Bundesverwaltung" veröffentlicht [15]. In Artikel 1 des Entwurfs findet sich eine Neufassung des BSI-Gesetzes wieder. Die explizite Verpflichtung zum Einsatz von Systemen zur Angriffserkennung findet sich nunmehr in § 31 „Besondere Anforderungen an die Risikomanagementmaßnahmen von Betreibern kritischer Anlagen" Abs. 2 wieder:

> „Betreiber kritischer Anlagen sind verpflichtet, Systeme zur Angriffserkennung einzusetzen. Die eingesetzten Systeme zur Angriffserkennung müssen geeignete Parameter und Merkmale aus dem laufenden Betrieb kontinuierlich und automatisch erfassen und auswerten. Sie sollten dazu in der Lage sein, fortwährend Bedrohungen zu identifizieren und zu vermeiden sowie für eingetretene Störungen geeignete Beseitigungsmaßnahmen vorzusehen. Dabei soll der Stand der Technik eingehalten werden. Der hierfür erforderliche Aufwand soll nicht außer Verhältnis zu den Folgen eines Ausfalls oder einer Beeinträchtigung der betroffenen kritischen Anlage stehen."

Diese Formulierung weicht nur in der Einleitung von der aktuell gültigen Formulierung aus dem IT-SiG 2.0 („Die Verpflichtung nach Absatz 1 Satz 1, angemessene organisatorische und technische Vorkehrungen zu treffen, umfasst ab dem 1. Mai 2023 auch den Einsatz von Systemen zur Angriffserkennung") ab und wurde insofern faktisch unverändert übernommen.

Im Zusammenspiel der beiden Gesetzgebungsprozesse sind die Verpflichtetenkreise „Besonders wichtige Einrichtungen", „Wichtige Einrichtungen" und Betreiber von „Kritischen Anlagen" konkret zu prüfen. Stand Juni 2024 bezieht sich die SzA-Verpflichtung nur auf Betreiber von Kritischen Anlagen, deren Bestimmung aber auch im Sektor Gesundheitswesen durchgeführt werden soll.

2.3.4 Orientierungshilfe Systeme zur Angriffserkennung

Das BSI hat zur Konkretisierung der gesetzlichen Anforderungen zu SzA im September 2022 die „Orientierungshilfe zum Einsatz von Systemen zur Angriffserkennung" [3] (OH SzA) bereitgestellt. Praktische Relevanz entfalten die Abschnitte 3 („Anforderungen") und 4 („Nachweis").

2.3.4.1 Anforderungen

Die Anforderungen sind in den drei Bereichen Protokollierung, Detektion und Reaktion formuliert. Innerhalb dieser drei Bereiche werden jeweils weitere Untergliederungen vorgenommen. Das Projektteam hat die Formulierungen auf Einzelanforderungen heruntergebrochen; letztendlich haben sich aus der OH SzA 65 Einzelanforderungen ergeben. Dabei fußen die Anforderungen maßgeblich auf Bausteinen des BSI IT-Grundschutz:

- Basisanforderungen von OPS.1.1.5 Protokollierung [16]
- Basisanforderungen von DER.1 Detektion von sicherheitsrelevanten Ereignissen [17]
- Basisanforderungen und Standardanforderungen von DER.2.1 Behandlung von Sicherheitsvorfällen [18]

2.3.4.2 Nachweis

Abschnitt 4 „Nachweis" der OH SzA definiert ein sogenanntes Umsetzungsgradmodell (Reifegradmodell im weiteren Sinne) und legt mit dem Unterabschnitt „Nachweiserbringung" fest, welche Anteile der Anforderungen bis zu welchem Datum umgesetzt werden müssen:

- Grundsätzlich sollte ein Umsetzungsgrad der Stufe 4 erreicht werden; demnach sind alle MUSS-Anforderungen für alle Bereiche erfüllt. Eine Nichtumsetzung von SOLLTE-Anforderungen muss stichhaltig und nachvollziehbar begründet sein. Zudem muss ein kontinuierlicher Verbesserungsprozess etabliert sein.
- Im ersten Nachweiszyklus wird vom BSI ein Umsetzungsgrad der Stufe 3 akzeptiert. MUSS-Anforderungen müssen ebenfalls bereits für alle Bereiche erfüllt sein. Im Gegensatz zur Stufe 4 reicht bei SOLLTE-Anforderungen jedoch die Prüfung auf Notwendigkeit und Umsetzbarkeit, deren Verbindlichkeit durch die Vokabel „idealerweise" noch einmal entkräftet wird. Ein kontinuierlicher Verbesserungsprozess muss nur in Planung sein.

Kombiniert man die Aussagen aus dem Abschnitt „Nachweis" mit jenen aus dem Gesetz zu den Nachweiszyklen („Betreiber Kritischer Infrastrukturen haben die Erfüllung der Anforderungen [...] alle zwei Jahre dem Bundesamt nachzuweisen."), so ist faktisch von Mai 2023 bis April 2025 die Erreichung von Umsetzungsgrad 3 nachzuweisen – im Anschluss ist immer Umsetzungsgrad 4 nachzuweisen.

2.4 Status quo branchenspezifische Sicherheitsstandards (B3S)

Das BSI pflegt eine Übersicht mit als geeignet festgestellten B3S[8]; demnach existieren Stand Juni 2024 13 B3S.

[8] https://www.bsi.bund.de/DE/Themen/KRITIS-und-regulierte-Unternehmen/Kritische-Infrastrukturen/Allgemeine-Infos-zu-KRITIS/Stand-der-Technik-umsetzen/Uebersicht-der-B3S/uebersicht-der-b3s_node.html

Krankenhäuser haben von der Möglichkeit B3S ebenfalls Gebrauch gemacht. Die aktuelle Version 1.2 des Branchenspezifischer Sicherheitsstandard „Medizinische Versorgung" [4] (B3S MV) datiert auf Dezember 2022 – die Eignungsfeststellung läuft gemäß BSI-Webseite Januar 2025 ab.

2.4.1 SzA im B3S Medizinische Versorgung 1.2

SzA sind bereits im B3S MV 1.2 abgebildet; die Eignung wurde vom BSI im Jan. 2023 festgestellt und ist auf zwei Jahre ab Bekanntgabe begrenzt. Sie bezieht sich auf die Anforderungen nach § 8a Abs. 1 und explizit auch auf die Anforderungen nach § 8a Abs. 1a. Primär ist das Thema in dem B3S folgendermaßen abgebildet:

Tabelle 1: Abbildung SzA in B3S MV 1.2

Thema	Controls
Vorfallerkennung und Behandlung (Abschnitt 6.9)	ANF-0076 – ANF-0084
Externe Informationsversorgung und Unterstützung (Abschnitt 6.11)	Insb. ANF-0091
Intrusion Detection / Prevention (Abschnitt 6.13.5)	ANF-0113 – ANF-0116
Beschaffungsprozesse (Abschnitt 6.13.14)	Insb. ANF-0148 und INF-0150
Protokollierung (Abschnitt 6.13.15)	ANF-0153 – ANF-0159

Weitere Anforderungen und Regelungen aus dem B3S tragen ebenfalls zur Erfüllung der Anforderungen aus der OH SzA bei; dazu gehören beispielsweise die Ausführungen aus Abschnitt 5.2 zur Implementierung eines Informations-Risikomanagements.

2.5 Aufschlüsselung „Stand der Technik"

Sowohl die IT-Sicherheitsgesetze selbst als auch die untergesetzlichen Regelungen wie insbesondere die Rechtsverordnungen aber auch die Handreichungen des BSI (insb. im Rahmen von Orientierungshilfen) beziehen sich konsequent auf den „Stand der Technik".
Insofern wurde im ersten Schritt aufgearbeitet, was „Stand der Technik" bedeutet und wie es in die IT-Sicherheitsgesetzgebung Deutschlands abgebildet ist.

2.5.1 BSI-Gesetz und Stand der Technik

Das IT-SiG 2.0 referenziert auf die „angemessenen organisatorischen und technischen Vorkehrungen" aus dem bestehenden IT-Sicherheitsgesetz und erweitert diese um Auflagen zur Angriffserkennung; bzgl. der Ausgestaltung werden einige Auflagen gegeben [1]. Konkret heißt es:

> ‚13. § 8a wird wie folgt geändert:
> [...]
> b) Nach Absatz 1 wird folgender Absatz 1a eingefügt:
> „(1a) Die Verpflichtung nach Absatz 1 Satz 1, angemessene organisatorische und technische Vorkehrungen zu treffen, umfasst ab dem 1. Mai 2023 auch den Einsatz von Systemen zur

> Angriffserkennung. Die eingesetzten Systeme zur Angriffserkennung müssen geeignete Parameter und Merkmale aus dem laufenden Betrieb kontinuierlich und automatisch erfassen und auswerten. Sie sollten dazu in der Lage sein, fortwährend Bedrohungen zu identifizieren und zu vermeiden sowie für eingetretene Störungen geeignete Beseitigungsmaßnahmen vorzusehen." '

Im ursprünglichen (ersten) IT-Sicherheitsgesetz [2] selbst erfolgte lediglich eine Referenzierung auf den „Stand der Technik":

> „Nach § 8 werden die folgenden §§ 8a bis 8d eingefügt:
>
> „§ 8a
>
> Sicherheit in der Informationstechnik Kritischer Infrastrukturen
>
> (1) Betreiber Kritischer Infrastrukturen sind verpflichtet, spätestens zwei Jahre nach Inkrafttreten der Rechtsverordnung nach § 10 Absatz 1 angemessene organisatorische und technische Vorkehrungen zur Vermeidung von Störungen der Verfügbarkeit, Integrität, Authentizität und Vertraulichkeit ihrer informationstechnischen Systeme, Komponenten oder Prozesse zu treffen, die für die Funktionsfähigkeit der von ihnen betriebenen Kritischen Infrastrukturen maßgeblich sind. Dabei soll der Stand der Technik eingehalten werden."

Insofern ergibt sich aus der Regelungskette, dass auch für Systeme zur Angriffserkennung Stand der Technik einzuhalten ist.

Eine weitere Unterfütterung erfolgte im Entwurf des ersten IT-Sicherheitsgesetzes der Bundesregierung vom 25.02.2015 im Rahmen der Gesetzesbegründung, welche an sich keine regelnde Wirkung entfaltet, jedoch Einblick in die thematische Einordnung seitens des Regelungsvorschlägers ermöglicht [19]:

> „Auf Grund der weitreichenden gesellschaftlichen Auswirkungen ist bei den technischen und organisatorischen Vorkehrungen der Stand der Technik zu berücksichtigen. Stand der Technik in diesem Sinne ist der Entwicklungsstand fortschrittlicher Verfahren, Einrichtungen oder Betriebsweisen, der die praktische Eignung einer Maßnahme zum Schutz der Funktionsfähigkeit von informationstechnischen Systemen, Komponenten oder Prozessen gegen Beeinträchtigungen der Verfügbarkeit, Integrität, Authentizität und Vertraulichkeit gesichert erscheinen lässt. Bei der Bestimmung des Standes der Technik sind insbesondere einschlägige internationale, europäische und nationale Normen und Standards heranzuziehen, aber auch vergleichbare Verfahren, Einrichtungen und Betriebsweisen, die mit Erfolg in der Praxis erprobt wurden. Die Verpflichtung zur Berücksichtigung des Stands der Technik schließt die Möglichkeit zum Einsatz solcher Vorkehrungen nicht aus, die einen ebenso effektiven Schutz wie die anerkannten Vorkehrungen nach dem Stand der Technik bieten."

Die weitere Konkretisierung dieses Standes der Technik wurde den Betreibern überlassen. Der Gesetzgeber hat jedoch eine Möglichkeit eingeräumt, dass nicht jeder einzelne Betreiber erneut diese

Auslegung durchführen und nachweisen muss. Dazu wurde in §8a folgende Regelung zur Option von B3S getroffen:

> „§ 8a
> Sicherheit in der Informationstechnik Kritischer Infrastrukturen
> [...]
> (2) Betreiber Kritischer Infrastrukturen und ihre Branchenverbände können branchenspezifische Sicherheitsstandards zur Gewährleistung der Anforderungen nach Absatz 1 vorschlagen. Das Bundesamt stellt auf Antrag fest, ob diese geeignet sind, die Anforderungen nach Absatz 1 zu gewährleisten."

Um den Betreiberorganisationen und ihren Verbänden Handlungssicherheit hinsichtlich der Eignung von B3S zu geben, werden vom BSI Orientierungshilfen (OH) herausgegeben. Eine Verbindlichkeit dieser OH ist konkret in der IT-Sicherheitsgesetzgebung nicht verankert; da das BSI diese Orientierungshilfen jedoch für die Prüf- und Eignungsfeststellungsprozesse anwendet, entfalten sie faktisch dennoch einen gewissen Grad an Verbindlichkeit.

2.5.2 Allgemein

In der deutschen Rechtssetzung kommt bei Generalklauseln für Bezugnahme auf technische Regelungen die sogenannte Drei-Stufen-Theorie zum Einsatz. Gemäß Handbuch der Rechtsförmlichkeit des BMJV [20] sollen nur die drei Stufen 1) allgemein anerkannte Regeln der Technik, 2) Stand der Technik und 3) Stand von Wissenschaft und Technik zur Anwendung kommen.
Die Auswahl ist entscheidend und richtet sich nach dem Gefährdungspotential der Regelungsmaterie. „*Stand der Technik*" befindet sich mit den Anforderungen zwischen den beiden anderen genannten Verweismöglichkeiten und ist gemäß Handbuch [20] (Abschnitt 4.5.1) wie folgt definiert:

> „Das Anforderungsniveau bei der Generalklausel „Stand der Technik" liegt zwischen dem Anforderungsniveau der Generalklausel „allgemein anerkannte Regeln der Technik" und dem Anforderungsniveau der Generalklausel „Stand von Wissenschaft und Technik". Stand der Technik ist der Entwicklungsstand fortschrittlicher Verfahren, Einrichtungen und Betriebsweisen, der nach herrschender Auffassung führender Fachleute das Erreichen des gesetzlich vorgegebenen Zieles gesichert erscheinen lässt. Verfahren, Einrichtungen und Betriebsweisen oder vergleichbare Verfahren, Einrichtungen und Betriebsweisen müssen sich in der Praxis bewährt haben oder sollten – wenn dies noch nicht der Fall ist – möglichst im Betrieb mit Erfolg erprobt worden sein."

2.5.3 Sonstige Ausgestaltung Stand der Technik

In Deutschland ist bezüglich der Ausgestaltung von „Stand der Technik" bezogen auf die IT-Sicherheit der „Bundesverband IT-Sicherheit e.V." (TeleTrust)[9] mit seinem Arbeitskreis[10] und einer entsprechenden Handreichung zum Thema [21] hervorzuheben.

In der Handreichung sind Erkennungstechnologien und wirksame Hilfsmittel wie IDS (Abschnitt 3.2.17), Endpoint Detection and Response (Abschnitt 3.2.22), Angriffserkennung und Auswertung (SIEM) (Abschnitt 3.2.24) sowie Cyber Threat Intelligence (Abschnitt 3.2.27) bereits abgebildet.

Eine Konkretisierung hinsichtlich der Anwendung in der Gesundheitsversorgung findet in diesem Zusammenhang nicht statt.

[9] https://www.teletrust.de
[10] https://www.teletrust.de/arbeitsgremien/ak-stand-der-technik/

3 Betreiberbefragung Umsetzung SzA

3.1 Ziel und Methodik

Das Hauptziel dieser Erhebung war die Erfassung des Ist-Zustands der Angriffserkennung in deutschen Krankenhäusern. Dies umfasste die Untersuchung der vorliegenden Netzwerkstrukturen, Sicherheitspraktiken und -vorkehrungen in diesen Einrichtungen, um potenzielle Herausforderungen beim Aufbau einer Angriffserkennungsinfrastruktur in der stationären medizinischen Versorgung zu identifizieren.

Die Datenerhebung und -analyse wurde mithilfe eines Mixed-Methods-Ansatzes durchgeführt, um eine umfassende Analyse zu ermöglichen. Um grundlegende Informationen über die befragten Krankenhausbetreiber zu erfassen, wurden Methoden der deskriptiven Statistik verwendet. Dazu gehören Durchschnittswerte, Häufigkeiten und Prozentsätze. Die Auswertung dieser statistischen Daten ermöglichte es, einen Überblick über grundlegende Merkmale der teilnehmenden Krankenhäuser sowie die Verteilung der Antworten auf geschlossene Fragen zu gewinnen.
Die Antworten auf offene Fragen in der Befragung, bei denen die Teilnehmer die Möglichkeit hatten, ihre Ansichten und Erfahrungen ausführlich darzulegen, wurden mittels qualitativer Analyse systematisch interpretiert, um Muster, Themen und Einsichten hinsichtlich der Angriffserkennung und der Herausforderungen bei der Sicherheitspraxis in Krankenhäusern zu identifizieren.
Durch die Kombination von deskriptiver Statistik und qualitativer Analyse konnten wir ein umfassendes Bild des Ist-Zustands von Angriffserkennung in deutschen Krankenhäusern zeichnen und wertvolle Einblicke in die Sicherheitspraktiken und -vorkehrungen in dieser Branche erhalten.

3.2 Rahmenbedingungen und Teilnehmerprofil

Die Datenerhebung erstreckte sich über den Zeitraum vom 3. Juli bis zum 18. August 2023. Über die Verteiler der Deutschen Krankenhausgesellschaft (DKG) und der Landeskrankenhausgesellschaften wurden nahezu alle Krankenhäuser in Deutschland um Teilnahme an der Befragung gebeten. Der Fragebogen wurde insgesamt 59-mal ausgefüllt, und die Befragten stammten aus einem breiten Spektrum von Krankenhäusern in Deutschland. Die Untersuchung deckte Krankenhäuser unterschiedlicher Größe ab, von kleinen Einrichtungen mit etwa 200 Mitarbeitenden und weniger als 200 Betten bis hin zu den größten Krankenhäusern mit etwa 18.000 Mitarbeitenden und über 800 Betten. Der Durchschnitt der befragten Krankenhäuser hatte etwa 4.276 Mitarbeitende, wobei der Median bei 2.500 Mitarbeitenden lag. Einige der Befragten waren für die Beantwortung des Fragebogens für mehrere Standorte verantwortlich (die höchste Anzahl von Standorten betrug 16), was zu einer Gesamtzahl von 132 abgedeckten Standorten führte.
Es ist wichtig anzumerken, dass die Beantwortung einzelner Fragen freiwillig war, Häuser sich also bei einzelnen Fragen enthalten konnten. Diese Möglichkeit wurde vor allem gegeben, um die Rücklaufquote

zu erhöhen. Aus Gründen der Transparenz wird die Anzahl der erhaltenen Antworten daher stets angegeben.

Nach Abschluss der Befragung fand am 28. August 2023 ein Auswertungs-Workshop statt. Dieser Workshop wurde im Rahmen des Unterarbeitskreises „B3S-Fortschreibung" des UP KRITIS Branchenarbeitskreises für die medizinische Versorgung durchgeführt.
Während des Workshops wurden den Branchenexperten die Ergebnisse der Befragung präsentiert, und es bot sich die Gelegenheit, die Ergebnisse kritisch zu hinterfragen, zu erläutern und einzuordnen. Darüber hinaus wurde in diesem Rahmen die branchenspezifischen Herausforderungen diskutiert, und die ersten Grundpfeiler für Empfehlungen zur Umsetzung der gesetzlichen Anforderungen gelegt. Dieser Workshop trug zur weiteren Validierung der Ergebnisse bei und förderte den fachlichen Austausch über Sicherheitspraktiken in Krankenhäusern.

3.3 Ergebnisse der Befragung

Im Folgenden werden die Ergebnisse der Befragung dargestellt, wobei besonderes Augenmerk auf die Reifegrade der SzA, die Risikolage und branchenspezifischen Gefährdungen gelegt wird. Zudem werden die Grundlagen der SzA sowie Detailfragen zu Protokollierung, Detektion und Reaktion, und der Datenschutz im Zusammenhang mit SzA behandelt. Der für die Befragung zugrundeliegende Fragebogen kann in Anlage 2: Fragebogen Bestandsaufnahme Umsetzung SzA eingesehen werden. Wenn es Unterschiede in den Antworten von KRITIS und nicht-KRITIS Betreibern gibt, werden diese hervorgehoben.

3.3.1 Selbsteinschätzung zum Reifegrad der SzA

Die Umfrageergebnisse zur Einschätzung des Reifegrades der Angriffserkennung (Abbildung 4) zeigen, dass sich die Mehrheit der befragten Häuser (75%) zum Zeitpunkt der Befragung bei einem Reifegrad kleiner als drei einordnet. Lediglich zwei KRITIS-Häuser geben an, Reifegrad vier erreicht zu haben. Das bedeutet, dass ein Viertel aller befragten Häuser die Anforderungen der OH-SzA zum Zeitpunkt der Befragung nach eigener Einschätzung erfüllt hat (n=57). Die KRITIS-Häuser ordnen sich mit 28% leicht darüber ein (n=40).

Abbildung 4: Selbsteinschätzung der Befragten zum Reifegrad SzA in ihrer Organisation

Betrachtet man die Umfrageergebnisse bezüglich ihrer Unterteilung in Informationstechnik (IT), Medizintechnik (MT) und Versorgungstechnik (VT) (Abb. 6), ist zu erkennen, dass die Angriffserkennung in der IT am weitesten fortgeschritten ist. Hier geben 26% der Befragten an, einen Reifegrad von mindestens drei erreicht zu haben (n=57). Unter den KRITIS-Häusern beläuft sich diese Zahl auf 30% (n=40). Die Medizintechnik folgt, 19% der Befragten geben an, Reifegrad drei der OH-SzA erreicht zu haben (n=57). Bei den KRITIS-Häusern sind es 25% (n=40). Es ist jedoch zu beachten, dass im Gegensatz zur IT 16% der Häuser angeben, bei der Medizintechnik noch keine Maßnahmen oder Planungen für die Angriffserkennung vorgenommen zu haben (KRITIS: 15%). Im Bereich der Versorgungstechnik (VT) geben insgesamt 16% der befragten Häuser an, einen Reifegrad von drei erreicht zu haben, wobei KRITIS-Häuser diesen Reifegrad etwas häufiger (20%) erreichen. Auffällig ist, dass 25% der Befragten angeben, weder Maßnahmen noch Planungen im Bereich der VT vorgenommen zu haben.

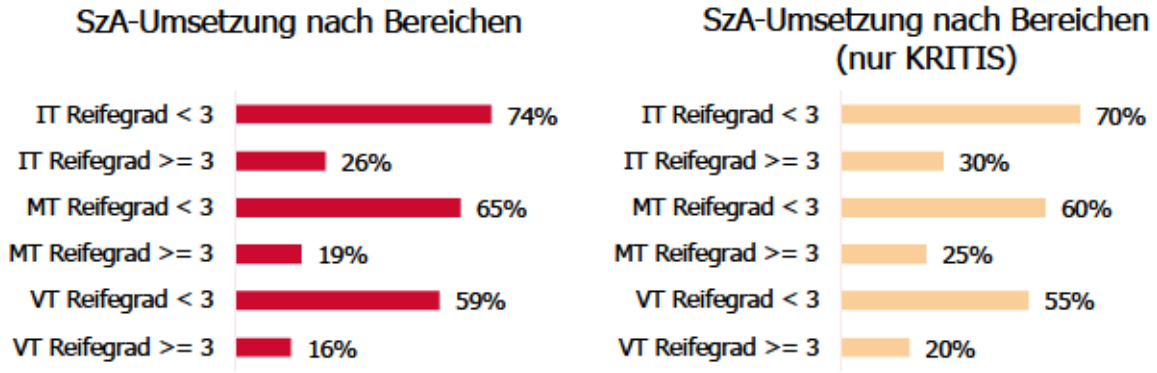

Abbildung 5: Selbsteinschätzung der Befragten zum Umsetzungsgrad SzA in IT, MT & VT

3.3.2 Risikolage und branchenspezifische Gefährdungen

In Bezug auf die Risikolage und branchenspezifischen Gefährdungen für deutsche Krankenhäuser (n=59), konnte festgestellt werden, dass es keinen signifikanten Unterschied zwischen den Antworten von KRITIS- und nicht-KRITIS-Häusern gab. Die identifizierten Top-Risiken für deutsche Krankenhäuser sind:

1. Nichtverfügbarkeit von behandlungsprozessrelevanten IT-Systemen (76%)
2. Nichtverfügbarkeit relevanter Daten im Diagnoseprozess (56%) und im Therapieprozess (49%)
3. Unterbrechung von behandlungsrelevanten Kommunikationsabläufen (44%)

Die Top-Bedrohungen für deutsche Krankenhäuser nach Einschätzung der Befragten (n=59, KRITIS n=41) sind:

1. Schadprogramme, insbesondere Ransomware (KRITIS und nicht-KRITIS: 90%)
2. Hacking und Manipulation (81%, KRITIS 90%)
3. Abhängigkeiten von Dienstleistern und Herstellern (KRITIS und nicht-KRITIS: 69%)

Erwähnenswert ist, dass Innentäter und unbefugte Zugriffe ebenfalls im Fokus der Krankenhäuser liegen, da 64% (KRITIS 71%) beabsichtigen, Systeme zur Angriffserkennung einzusetzen, um dieser Bedrohung zu begegnen. Außerdem ist im Zusammenhang mit den Top-Bedrohungen auffällig, dass sich die erste Bedrohung für Krankenhäuser, die nicht oder nicht direkt durch SzA adressiert werden kann, auf Platz 5 befindet (Ausfall von Basisinfrastrukturen mit direktem Bezug zu IT (Sekundäreffekte, z. B. Strom und TK, 61%). Im Rahmen des Auswertungs-Workshops wurde von den Experten unterstrichen, dass dies eine besondere Herausforderung mit hohem Schadenspotential darstellt. Es lässt sich aus den gegebenen Antworten daher schließen, dass derzeit das Bewusstsein für Bedrohungen durch Cyber-Angriffe präsent ist.

Bedrohungsinformationen (n=59, KRITIS n=41) beziehen die Befragten nach eigener Angabe hauptsächlich von Herstellern von Hard- und Software (92%, KRITIS 90%), von Behörden, insbesondere dem Bundesamt für Sicherheit in der Informationstechnik (BSI) (81%), sowie aus verschiedenen (Fach-)Medien, wie beispielsweise Heise oder Golem (78%).

3.3.3 Grundlagen für den Aufbau von SzA

Als Good Practices für IT-Sicherheit zogen zum Zeitpunkt der Befragung 83% der Befragten den „B3S" heran (n=59). In der KRITIS-Stichprobe liegt dieser Anteil bei 85%, wodurch die hohe Relevanz dieses Dokuments für die Branche unterstrichen wird (n=41).

Für die Umsetzung von SzA stützen sich 73% der Befragten (n=59) auf die OH-SzA des Bundesamts für Sicherheit in der Informationstechnik (BSI). In der KRITIS-Stichprobe sind es 78% (n=41). Darüber hinaus wird der Mindeststandard zur Protokollierung und Detektion von Cyber-Angriffen von 59% der

Befragten herangezogen. Dies zeigt, dass die Einhaltung von Standards und Richtlinien eine wichtige Rolle bei der Umsetzung von Sicherheitsmaßnahmen in deutschen Krankenhäusern spielt.

In Bezug auf die Nutzung von Frameworks zum Aufbau und Betrieb von SzA (z. B. MITRE ATT&CK) zeigt die Umfrage, dass etwa ein Drittel der befragten Einrichtungen bereits solche Frameworks verwenden, während weitere 46% planen, diese in Zukunft zu nutzen (n=59). Die angegebene Motivation für den Einsatz war vielfältig, wodurch die Potentiale solcher Frameworks für die Branche unterstrichen wird. Diese Erkenntnisse weisen auf ein gesteigertes Interesse und einen wachsenden Bedarf hin, die Sicherheitsmaßnahmen in Krankenhäusern zu strukturieren und zu optimieren.

Bei der Betrachtung der von den Befragten betriebenen Geräte gibt es eine große Spannbreite zwischen den Häusern. Im Durchschnitt werden 7.838 (Median: 1.400) Medizingeräte pro Haus betrieben, wobei das Maximum bei 60.000 und das Minimum bei 10 Geräten liegt. Netzwerkfähige Medizingeräte sind ebenso ubiquitär, zwischen 10 und 20.000 solcher Geräte befinden sich pro Krankenhaus im Einsatz, der Durchschnitt liegt bei 2.178 (Median: 500). Die Anzahl der IT-Geräte pro Krankenhaus liegt zwischen 300 und 120.000, mit einem Durchschnitt von 15.832 (Median: 4.750). Im Bereich der Versorgungstechnik liegt die Spanne der im Einsatz befindlichen Systeme bei 15 bis 50.000 Geräten, mit einem Durchschnitt von 2.258 (Median: 475).

Die Strukturanalyse der Netzsegmentierung (Abbildung 6) in den Krankenhäusern zeigt, dass 30% der Befragten (n=56) keine Netzsegmentierung verwenden oder diese bisher nicht durchsetzen (z. B., weil sie nur für strukturgebende und ordnende Zwecke etabliert wurden). Weitere 16% setzen sie lediglich zur Sicherung einzelner Geräte ein. Bei KRITIS-Häusern haben 26% keine durchgesetzte Netzsegmentierung, 21% setzen Regeln ausschließlich für einzelne Geräte durch (n=40). Hinsichtlich der Anzahl der VLANs reicht die Spanne von 3 bis 1300, wobei der Durchschnitt bei 151 liegt. Gerade im Verhältnis zu den durchschnittlich im Betrieb befindlichen, netzwerkfähigen Geräten lässt dies auf sehr große Netzsegmente schließen. Die Gruppierung der Geräte erfolgt insbesondere nach der Unterscheidung zwischen Medizintechnik, Informationstechnik und Versorgungstechnik. 64% gaben dies als maßgebliches Kriterium für die Segmentzuordnung an (n=59), bei KRITIS-Häusern waren es 71% (n=41).

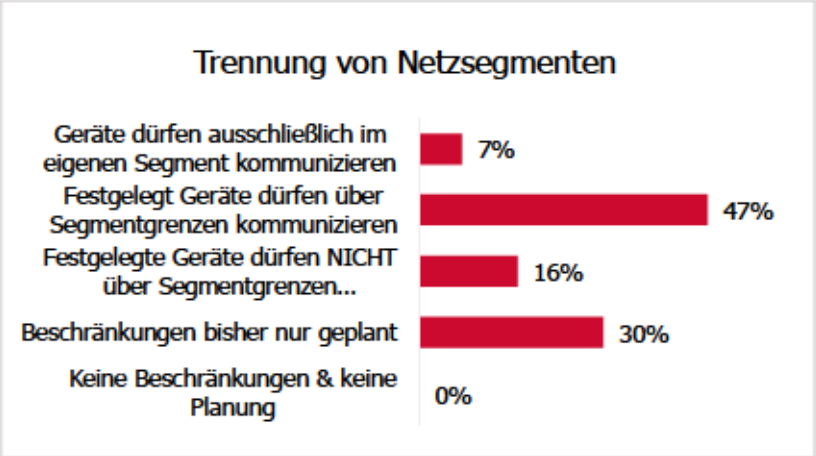

Abbildung 6: Die Durchsetzung von Netzsegmentgrenzen

Erwähnenswert ist zudem eine Divergenz beim Einsatz von Cloudtechnologie. 32% der Befragten (n=44) geben an, Clouddienste zu nutzen, wobei in der KRITIS-Stichprobe dieser Anteil bei 38% liegt (n=34). Andere machen insbesondere in Freitextantworten ihre Bedenken deutlich und lehnen den Einsatz pauschal ab.

Als Kriterium für die Reihenfolge, in der die Systeme und Netzwerkkomponenten in die Angriffserkennung aufgenommen wurden oder werden sollen, gaben 68% der Befragten die Kritikalität der Systeme an, während 39% dies nach der Relevanz der Systeme für die kritische Dienstleistung (kDL) tun (n=59).

3.3.4 Protokollierung

Die Umfrage zur Protokollierung von Daten zeigt, dass 56% der befragten Einrichtungen Protokolldaten zum Zeitpunkt der Erfassung ausschließlich lokal vorhalten. Hingegen setzen 44% bereits eine Form der zentralen Protokollierung ein. Bei der zentralen Protokollierung sind im Durchschnitt bereits 1.144 Systeme eingebunden, wobei die Spanne von 3 bis 20.000 reicht, und der Median bei 100 liegt (n=26). Dies legt nahe, dass einzelne Krankenhäuser bereits weiter fortgeschritten sind, es aber einen großen Teil von Krankenhäusern gibt, die noch nicht sehr viele Systeme in die zentrale Protokollierung aufgenommen haben.

In Bezug auf die Protokollierung im Bereich der Medizintechnik zeigte die Umfrage, dass den meisten Befragten keine klare Aussage darüber möglich war, ob netzwerkfähige Medizingeräte in der Lage sind, Logdaten zu versenden. 52% der Befragten (n=48) und 53% der nicht-KRITIS-Häuser (n=34) wählten die Antwortoption „unbekannt". 21% der Befragten gaben an, dass weniger als 10% der Geräte dazu in der Lage wären, wobei dieser Anteil bei KRITIS-Häusern bei 18% lag.

Die Antworten zur Protokollierung von Netzverkehrsdaten zeigen: Bei 60% der Befragten (n=50) erfolgt die Protokollierung an den Netzübergängen von intern zu extern (56% in der KRITIS-Stichprobe, n=36). 22% protokollieren Verkehr an Netzübergängen von intern zu intern (n=50). Innerhalb von internen Netzbereichen protokollieren 15% der Befragten Netzwerkverkehr (n=47), wiederum ohne Unterschied zwischen KRITIS- und nicht-KRITIS-Häusern. Insgesamt verdeutlichen diese Ergebnisse die Vielfalt der Protokollierungspraktiken und den Einsatz von Angriffserkennungstechnologien in deutschen Krankenhäusern.

3.3.5 Detektion

In Bezug auf die Detektion von Sicherheitsvorfällen setzen 83% der befragten Einrichtungen (n=59) Next-Generation Firewalls ein, wobei dieser Anteil in KRITIS-Häusern bei 78% (n=41) liegt. Diese Systeme übernehmen in diesen Fällen auch die Auswertung der Netzverkehrsdaten (81% in der

Gesamtstichprobe, 83% in KRITIS-Häusern). Zusätzlich ziehen 20% der Befragten Metadaten des Netzwerkverkehrs und 15% Network-Flow-Daten zur Analyse heran.

83% der Befragten verwenden eine Antivirus-Software mit zentralem Management, wobei dieser Wert bei KRITIS-Häusern bei 80% liegt. Des Weiteren setzen 85% der Befragten Mail-Filtering-Technologien ein.

11% der Einrichtungen haben einen Managed Security Service Provider (MSSP) für Teile der Angriffserkennung beauftragt, während weitere 76% den Einsatz prüfen oder sich bereits in der Vorbereitungsphase befinden. Dies zeigt ein wachsendes Interesse an Outsourcing-Lösungen in diesem Bereich.

Befragt nach den durch SzA abgedeckten Bereichen zum Zeitpunkt der Befragung ergeben sich die in Abbildung 7 gezeigten Zahlen. Auffällig ist, dass diese scheinbar der Selbsteinschätzung zum Reifegrad der SzA in diesen Bereichen in Kapitel 3.3.1 widersprechen. So gaben 74% an, sich im Bereich der IT in Reifegrad 1 oder 2 zu befinden (n=57), während nur 61% angeben, tatsächlich mit Aufbau und Betrieb von SzA in diesem Bereich begonnen zu haben (n=59). Eine genauere Untersuchung zeigt, dass 30% der Häuser mit Reifegrad eins zum Zeitpunkt der Befragung ausschließlich die Planung finalisiert, jedoch noch nicht mit der Umsetzung begonnen haben. Diese Daten verdeutlichen zudem, dass die Erreichung eines Reifegrades von mindestens drei der OH SzA (nach Selbsteinschätzung 19% MT und 16% VT) nicht für eine komplette Abdeckung dieser branchenspezifischen Bereiche genügt, da dies lediglich 3% im Bereich der Medizintechnik bzw. 2% im Bereich der Versorgungstechnik für sich reklamieren.

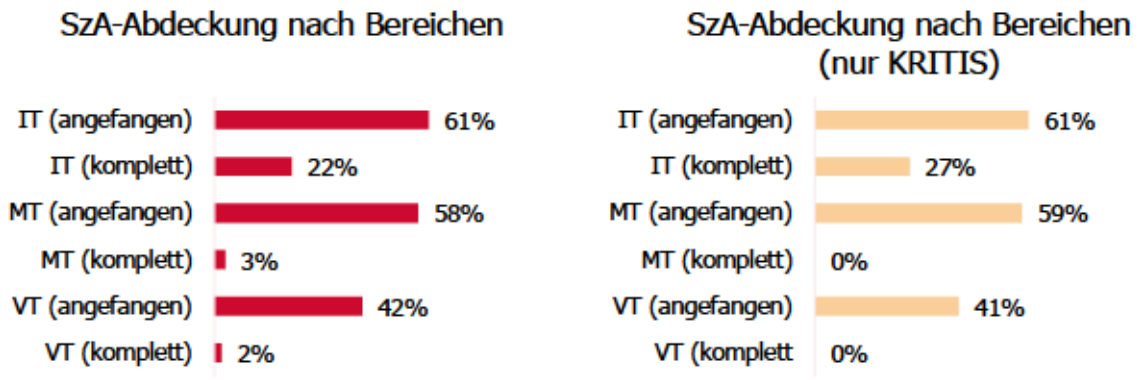

Abbildung 7: Durch SzA abgedeckte Bereiche zum Zeitpunkt der Befragung

Die Umfrageergebnisse zeigen, dass die Detektion in etwa der Hälfte der befragten Einrichtungen überwiegend automatisiert erfolgt. 42% geben dagegen an, dass die Erkennung primär manuell vorgenommen wird, wobei bei 24% die Erkennung ausschließlich manuell erfolgt.

Diese Ergebnisse verdeutlichen die Vielfalt der eingesetzten Detektionstechnologien und -praktiken in deutschen Krankenhäusern und zeigen gleichzeitig Entwicklungsmöglichkeiten auf, insbesondere im Bereich der automatisierten Erkennung und der Reaktionszeiten auf Sicherheitsvorfälle.

3.3.6 Reaktion

Für eine schnelle Reaktion auf IT-Sicherheitsvorfälle können Arbeitsabläufe im Voraus geplant werden. So genannte Playbooks unterstützen Organisationen dabei, für sicherheitsrelevante Szenarien, etwa die Kompromittierung eines Systems mit Ransomware, Maßnahmen festzulegen. Runbooks können darüber hinaus konkrete IT-Abläufe definieren, z. B. wie ein System aus einem Backup wiederhergestellt werden kann. 66% der befragten Einrichtungen (n=47) nutzen diese Möglichkeiten der Prozessverschriftlichung für IT-Sicherheitsvorfälle. 17% haben zum Zeitpunkt der Befragung mehr als fünf Play- oder Runbooks definiert. 34% geben hingegen an, keine Play- oder Runbooks zu nutzen. In KRITIS-Häusern haben 23% mehr als fünf Play- oder Runbooks und 37% keine (n=35).
Reaktionszeiten für Alarme haben zum Zeitpunkt der Befragung 46% der Organisationen definiert. In der konkreten Ausgestaltung der Zeiten zeigen sich Unterschiede. 22% geben an, dass sie nach einigen Stunden reagieren, während 14% eine Reaktion auf Alarme innerhalb von Minuten anstreben.

22% der Einrichtungen (n=50) verfügen über automatisierte Reaktionsmöglichkeiten, wobei dieser Anteil in KRITIS-Häusern bei 19% (n=37) lag. Die automatisierte Reaktion findet in 82% der Fälle in der Informationstechnik und in 40% in der Versorgungstechnik statt. In KRITIS-Häusern ist die automatisierte Reaktion in allen Fällen in der IT verortet, in 17% darüber hinaus auch in der VT.
Es ist wichtig festzustellen, dass keiner der Befragten automatisierte Reaktionen im Bereich der Medizintechnik einsetzt, was auf die besondere Sensibilität dieses Bereichs und die Notwendigkeit manueller Interventionen hinweisen könnte.

3.3.7 Datenschutz

Zum Datenschutz im Kontext von SzA in deutschen Krankenhäusern sind folgende Erkenntnisse aus der Analyse hervorzuheben:
29% der Befragten gaben an, das Security Information and Event Management System (SIEM) bereits in das Verzeichnis der Verfahrenstätigkeiten aufgenommen zu haben. Dies ist ein wichtiger Schritt, um sicherzustellen, dass der Umgang mit personenbezogenen Daten im Einklang mit den Datenschutzvorschriften erfolgt. In KRITIS-Häusern liegt dieser Anteil bei 36%. Die Protokollierung von Zugriffen auf das SIEM wird von 33% der Befragten durchgeführt, wobei in KRITIS-Häusern 39% diese Praxis verwenden. Eine beträchtliche Anzahl von Befragten gibt an, dass diese Informationen unbekannt sein, nämlich 38% in der Gesamtstichprobe und 44% in KRITIS-Häusern.
Die Speicherdauer der Rohdaten der Protokollierung zeigt eine erhebliche Varianz, von 7 Tagen bis zu mehreren Jahren oder sogar unbegrenzt. Dies unterstreicht die Bedeutung klarer Richtlinien und Prozesse zur Datenlöschung und -speicherung, um den Datenschutzbestimmungen gerecht zu werden und die Integrität personenbezogener Daten zu wahren.

Laut OH-SzA müssen Mitarbeitende für SzA namentlich benannt werden. Die Umfrage ergab, dass 32% der Einrichtungen keine Personen namentlich benannt hatten, wobei dieser Anteil in KRITIS-Häusern bei

16% liegt. Unter den Befragten, die Mitarbeitende benannt hatten, war die Zahl der Benannten sehr unterschiedlich: 19% der Einrichtungen hatten einen Mitarbeitenden benannt, 49% hatten 2-5 Mitarbeitende namentlich benannt, und in einer Organisation waren sogar 10 Personen namentlich benannt. Die Ergebnisse der Befragung zeigen die Notwendigkeit, auch den Datenschutz im Kontext des Aufbaus einer Detektionsinfrastruktur mitzudenken, um die sensiblen Daten, die bei der für die SzA notwendigen Protokollierung anfallen, zu schützen und nicht zuletzt, um die Einhaltung von geltenden Datenschutzvorschriften sicherzustellen.

3.4 Diskussion und Schlussfolgerungen

3.4.1 Limitationen

Die Gesamtzahl der Krankenhäuser in Deutschland lag dem statistischen Bundesamt zufolge im Jahr 2022 bei 1.893[11]. Die genaue Anzahl der KRITIS-Krankenhäuser in Deutschland ist nicht öffentlich bekannt, aber gemäß Daten aus der BSI-Studie zum KRITIS-Sektor Gesundheit aus dem Jahr 2020 hatten sich etwa 10 Prozent der deutschen Krankenhäuser beim Bundesamt für Sicherheit in der Informationstechnik (BSI) als Betreiber Kritischer Infrastrukturen registriert. [22] Wenn diese Maßzahl auf die Zeit der Befragung angewendet wird, dürfte es etwa 180 KRITIS-Krankenhäuser in Deutschland gegeben haben. Dieser Rechnung zufolge hätten etwa ein Viertel aller KRITIS-Häuser in Deutschland an der Befragung teilgenommen. Dies unterstreicht die Relevanz der Studienergebnisse für KRITIS-Einrichtungen.

Eine wesentliche Limitation dieser Erhebung besteht darin, dass deutlich mehr KRITIS-Häuser als nicht-KRITIS-Häuser an der Umfrage teilgenommen haben. Mehr als 70 Prozent (70,2%) der Antworten kamen von Krankenhäusern, die als Betreiber Kritischer Infrastrukturen eingestuft sind. Verschiedene Faktoren können dazu beigetragen haben, darunter das größere Bewusstsein für Sicherheitsfragen in KRITIS-Einrichtungen oder eine stärkere Motivation zur Teilnahme durch die zuvor erläuterten, neuen gesetzlichen Vorgaben, die zum Zeitpunkt der Befragung KRITIS-Häuser unmittelbar betraf. Jedoch ist zu beachten, dass KRITIS-Häuser im Gesamtkontext der deutschen Krankenhauslandschaft eine vergleichsweise kleinere Gruppe darstellen. Aus diesem Grund sind die Ergebnisse dieser Erhebung nicht zwangsläufig repräsentativ für die Gesamtheit der deutschen Krankenhäuser. Die überproportionale Beteiligung von KRITIS-Krankenhäusern kann zu Verzerrungen führen, da die spezifischen Herausforderungen und Sicherheitspraktiken in KRITIS-Einrichtungen möglicherweise nicht oder nicht vollständig auf andere Krankenhaustypen übertragbar sind.

Trotz dieser Limitationen ist es jedoch möglich, allgemeine Trends und Muster für Krankenhäuser aus den Ergebnissen herauszuarbeiten und Indikatoren über den Ist-Zustand von SzA in deutschen Krankenhäusern abzuleiten. Die Erhebung bietet wertvolle Einblicke in die Sicherheitspraktiken und

[11] https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Krankenhaeuser/Tabellen/gd-krankenhaeuser-jahre.html

Herausforderungen in dieser speziellen Gruppe von Krankenhäusern, die als Betreiber Kritischer Infrastrukturen eingestuft sind.

3.4.2 Schlussfolgerungen

In der Strukturanalyse zeigte sich, dass die digitalen Infrastrukturen deutscher Krankenhäuser sehr unterschiedlich sind. Neben offensichtlichen Asset-Mengenunterschieden zeigt sich dies beispielsweise in der Cloudnutzung, die von 68% abgelehnt wird, während 32% Cloudtechnologie in unterschiedlicher Intensität in ihre Infrastruktur eingebunden haben und nutzen. Ein weiteres Beispiel ist die Netzwerksegmentierung. Die Segmentstrukturierung erscheint sehr divers. Den Durchschnittswerten zufolge sind Segmente überwiegend sehr groß bzw. beinhalten sehr viele Assets pro Segment. Auch werden Geräte unterschiedlich auf Segmente aufgeteilt. Zwei Drittel der Befragten gaben an, dass die Aufteilung primär nach der Kategorisierung IT – MT – VT erfolgt sei. Ein Drittel wählte eine abweichende Zuordnung der Assets zu Netzwerksegmenten. Die qualitative Analyse ergab allerdings, dass insbesondere die Netzwerksegmentierung von einigen Häusern im Rahmen des Krankenhauszukunftsgesetzes (KHZG) zum Zeitpunkt der Befragung gerade erneuert wurde.
Die VT scheint nicht nur in Bezug auf SzA etwas weniger im Fokus zu stehen, sondern die Datenlage scheint insgesamt etwas geringer als bei IT und MT zu sein. Während es bei der Abfrage der anderen Assetklassen IT und MT 50 und 46 Antworten gab, sich also nur 9 bzw. 14 der Befragten enthielten, beantworteten die Frage zu VT-Assets lediglich 36 der Befragten. Hier enthielten sich also 23 Befragte. Dies könnte damit zusammenhängen, dass konkrete Zahlen nicht vorlagen. Diese Schlussfolgerung passt zu einer mehrfach in unterschiedlichen Kontexten benannten Herausforderung: dem Thema Observability.
Insgesamt lässt sich feststellen, dass es keine einheitliche Empfehlung für den Aufbau einer SzA-Infrastruktur geben wird, die allen Krankenhäusern und vorherrschenden Bedingungen gerecht wird.
Grundsätzlich zeigte sich in der Befragung aber, dass Krankenhäuser, egal ob KRITIS oder nicht-KRITIS, Bedrohungen und Risikolage sehr ähnlich einschätzten. Auch stehen die Häuser vor ähnlich gelagerten Herausforderungen, die bewältigt werden müssen. Insbesondere ist eine oft genannte Herausforderung, gute Quellen für CTI zu finden. Hier scheint es großen Bedarf nach einer qualitativen Austauschplattform zu geben.

Bei der Protokollierung zeigen sich vor allem Probleme in den branchenspezifischen Bereichen Medizin- und Versorgungstechnik. Medizintechnik bietet oft keine auskömmlichen Möglichkeiten, Logs an eine zentrale Instanz weiterzuleiten. Auch die wahrnehmbare und nachvollziehbare Vorsicht bei der Reaktion auf Angriffe im Bereich der Medizintechnik ist ein Thema, das die deutschen Krankenhausbetreiber vor eine große Herausforderung stellt. Insbesondere dieser Bereich unterscheidet die Branche von anderen und allgemeine Vorgaben greifen zu kurz. Dies wird noch einmal unterstrichen durch die Diskrepanz zwischen der Erreichung des Reifegrades drei nach OH SzA in den Bereichen der Medizin- und Versorgungstechnik und der nach Einschätzung der Befragten trotzdem nicht vollständig durch SzA

abgedeckten Bereiche. Es lässt sich schlussfolgern, dass die durch die OH-SzA empfohlenen, generellen Maßnahmen, die eine Orientierung für KRITIS-Betreiber aller Sektoren bieten sollte, durch branchenspezifische Empfehlungen erweitert werden müssen, damit eine umfassende Absicherung durch SzA gelingen kann.
Die Motivation zur Heranziehung von Frameworks wie MITRE ATT&CK zum Aufbau der Detektionsinfrastruktur zeigt die Bereitschaft der Befragten und ihre Suche nach verwendbaren Ressourcen. Die großen Unterschiede bei Einsatz und Umsetzung verdeutlichen aber auch, dass vorhandene Good Practices nicht einfach auf die Branche der medizinischen Versorgung zu übertragen sind. In der Konsequenz ergibt sich daher ein großer Bedarf nach auf die medizinische Versorgung zugeschnittenen Empfehlungen.

Die Reaktion scheint zum Zeitpunkt der Befragung noch nicht im Fokus des SzA-Aufbaus gestanden zu haben. Das Gros der Befragten gab an, wenige oder keine verschriftlichten Play- oder Runboooks zu haben. Automatisierte Reaktionen gab es lediglich im Bereich der IT und auch dort nur von 22% der Befragten.
Auch Reaktionszeiten waren überwiegend undefiniert. Bei den Befragten, bei denen es definierte Zeiten gab, variierten sie stark von einigen Stunden hin zu Minuten im 24/7 Betrieb.
Die Ergebnisse verdeutlichen, dass die Reaktionsfähigkeiten auf Sicherheitsvorfälle in deutschen Krankenhäusern variieren, weiterentwickelt und vereinheitlicht werden müssen. Eine verstärkte Dokumentation von Play- und Runbooks und die Implementierung automatisierter Reaktionen können dazu beitragen, die Sicherheit deutscher Krankenhäuser zu stärken.

Beim Datenschutz gaben 71% der Befragten an, dass zumindest teilweise Anwendungsdaten in die Auswertung einbezogen würden. Den Umgang mit Patientendaten in Logdaten führten die Befragten allerdings sehr unterschiedlich durch. Während die einen alle Logdaten als sensible Daten mit hohem Schutzbedarf i.S.d. DSGVO ansehen, geben andere an, dass sämtliche Zugriffe auf die Logdaten von der Geschäftsleitung genehmigt werden müssten und wieder andere, dass die Daten lediglich on-premise gespeichert würden und damit dem Datenschutz genüge getan sei. Zusätzlich gaben drei Viertel der Befragten an, dass Zugriffe auf das SIEM gar nicht protokolliert würden oder dass zumindest unbekannt sei, ob diese Zugriffe protokolliert würden. Auch bei dieser Frage gab es auffallend viele Enthaltungen. Obwohl es die Antwortmöglichkeit „unbekannt" gab, beantworteten nur 71% der Befragten die Frage. Bei der Dauer der Aufbewahrung von Protokolldaten gab es deutliche Variation unter den Antworten mit maximalen Werten von 10 Jahren und unbegrenzter Speicherung.
Insgesamt unterstreichen diese Ergebnisse die Notwendigkeit, den Datenschutz im Rahmen der Angriffsdetektion mitzudenken und einheitliche Empfehlungen zu geben. Die Integration des SIEM in das Verzeichnis der Verfahrenstätigkeiten, transparente Protokollierung von Zugriffen und klare Richtlinien zur Speicherdauer sind Schlüsselaspekte, um Datenschutzvorschriften zu erfüllen und die Sicherheit von Patientendaten zu gewährleisten.

3.5 Fazit Betreiberbefragung

In der Befragung wurde der Umsetzungsstand der Angriffserkennung in deutschen Krankenhäusern erfasst. Während sich durch die hohe Beteiligung von KRITIS-Krankenhäusern einige Limitationen insbesondere bezüglich der Aussagekraft der Ergebnisse auf kleinere Krankenhäuser ergeben, können dennoch einige Schlüsse gezogen und Trends abgeleitet werden.

Die Befragung zeigt, dass Krankenhäuser sehr unterschiedliche digitale Infrastrukturen betreiben und ebenfalls unterschiedlich weit fortgeschritten sind, was den Aufbau von SzA angeht. Es gibt einzelne Häuser, die bereits viele Systeme mittels SzA überwachen. Das Gros der Häuser war zum Zeitpunkt der Befragung noch in der Aufbauphase. Dies ergibt sich einerseits aus der Selbsteinschätzung zum Reifegrad und andererseits aus den erfassten Antworten zu Detailfragen. Grundsätzliche oder pauschale Empfehlungen für den Aufbau von SzA sind dadurch zwar schwierig, dennoch zeigen sich Gemeinsamkeiten sowohl bei Risiken und Bedrohungen als auch bei Hürden und Herausforderungen, die adressiert werden müssen. Zu den größten Herausforderungen, denen sich Krankenhäuser beim SzA-Aufbau stellen müssen, gehören branchenspezifische Bedrohungsinformationen, sowie Protokollierung und Reaktion im Kontext von branchenspezifischer Technik (insb. Medizintechnik). Nicht zuletzt muss auch der Datenschutz bei Aufbau und Betrieb von SzA mitgedacht werden. Branchenspezifische Good Practices werden nachgefragt und sind für eine branchenweite Adaptierung unerlässlich. Eine aktualisierte Fassung des B3S, der auf diese Spezifika eingeht, könnte hier einen enormen Mehrwert für die deutschen Krankenhäuser bieten.

4 Erarbeitung Handlungsempfehlungen

Die Erarbeitung von Handlungsempfehlungen erfolgte entlang der folgenden Struktur, welche sich auch in den Unterabschnitten hierzu widerspiegelt:

- Durchführung Experteninterviews
- Sichtung von SzA-relevanten nationalen und internationalen Dokumenten
- Ergebnisse der Analyse
- Ableitung von Schlussfolgerungen

4.1 Expertengespräche

Zur Vertiefung der Informationsbasis dahingehend, wie der Stand der Technik bezüglich IT-Sicherheit in der Branche der Krankenhäuser bisher definiert wurde, wie Systeme zur Angriffserkennung in der Praxis umgesetzt werden, wie der IT-Sicherheitsregulierer das Thema bewertet und wie andere Branchen (mit gewissen Parallelen bei der strukturellen Ausgestaltung) umgehen, wurden folgende 8 Gespräche zu den Fragestellungen durchgeführt.

Termin	Thematischer Schwerpunkt des Ansprechparters	Teilnehmende
12.01.2024	Arbeitsweise des Arbeitskreises Stand der Technik innerhalb des BAK Medizinische Versorgung	Dr. Stefan Bücken (UK Erlangen)
23.01.2024	Perspektive des BSI (als Regulierer für IT-Sicherheit in Deutschland) auf das Thema SzA in Krankenhäusern	Martin Apel (BSI) Bernhard Steffens (BSI)
23.01.2024	Umsetzung von SzA in einem großen Krankenhaus (KRITIS)	Peter Weidenbach (UK Bonn)
25.01.2024	Regelungssetzung zu und Umsetzung von SzA in einer anderen KRITIS-Branche mit hohem OT-Anteil	Christian Cichowski (Wupperverband)
26.01.2024	Historie zur Erarbeitung des B3S Medizinische Versorgung	Markus Holzbrecher-Morys (DKG e.V.)
29.01.2024	Branchenübergreifende Sicht auf das Thema SzA beim Betrieb Kritischer Infrastrukturen in Deutschland (Experten allesamt im Themenarbeitskreis Detektion des UP KRITIS aktiv)	Anders Kölligan (BSI) – Leiter TAK Robert Steffen (Vodafone) – Sprecher TAK Oliver Mora (BWB) Stephan Dambach (SWS)
31.01.2024	Regelungssetzung zu und Umsetzung von SzA in einer anderen KRITIS-Branche mit hohem OT-Anteil	Mathias Böswetter (BDEW) Vertreter Betreiberunternehmen
31.01.2024	Umsetzung von SzA in einem mittelgroßen Krankenhaus (kein KRITIS)	Karsten Sydow (UK Brandenburg)

Alle Expertengespräche wurden als semistrukturierte Interviews durchgeführt. Die insgesamt 17 Fragen waren jeweils in die folgenden 3 Themenbereiche einsortiert:

- Validierung von SdT-Empfehlungen
- Anwendbarkeit von SdT-Empfehlungen
- Erfahrungswerte bei der Umsetzung von IT-Sicherheit

Die Gesamtheit der Fragen ist diesem Dokument als „Anlage 2: Fragebogen Bestandsaufnahme Umsetzung SzA" beigefügt. Die Auswertung der Interviews ist gesamtheitlich über den vollständigen Prozess zur Erarbeitung von Handlungsempfehlungen in Abschnitt 4.3 dokumentiert.

4.2 Sichtung nationaler / internationaler Standards und Good Practices

Es wurde eine weitreichende Recherche von Standards und Vorgehensweisen zu Informationssicherheit und IT-Sicherheit durchgeführt. Die Benennung und Auswahl bedienten sich unterschiedlicher Quellen:

- Die Orientierungshilfe Systeme zur Angriffserkennung des BSI (OH SzA) [3] selbst referenziert Dokumente.
- ENISA hält für Betreiber Kritischer Infrastrukturen eine Übersicht[12] vor, bei welcher die relevanten Themenbereiche auf die zutreffenden Abschnitte / Controls in ausgewählten Standards verlinkt sind.
- Die Experten aus den Expertengesprächen (s.u. Abschnitt 4.1 für Details) haben Hinweise auf einschlägige Dokumente eingesteuert.
- Die MedSec-Gruppe[13] selbst hat ihr Hintergrundwissen ebenfalls in die Auswahl eingesteuert.

Die Auswertung eines jeden Dokumentes erfolgte nach folgendem Vorgehen:

- Überblicksartige Sichtung des Gesamtdokumentes auf inhaltliche Anteile mit SzA-Relevanz (Einleitung und Inhaltsverzeichnis)
- Konkrete Prüfung der Anteile mit SzA-Relevanz, hinsichtlich:
    - Konkretisierungsvorschläge für die Gesundheitsversorgung im Spezifischen
    - Auslegungsvorschläge SzA allgemein, die auf die Gesundheitsversorgung angewendet werden können.

Die gesichteten Dokumente sind nachfolgend untergliedert in die inhaltlichen Bereiche aufgeführt:

- Standards mit Bezug zu SzA: KI-Sektor-Übergreifend, d.h. nicht konkret auf die Gesundheitsversorgung bezogen,
- Good Practice zu IT-Sicherheit (und ggf. SzA) in der stationären Gesundheitsversorgung,
- Nationaler Input – Sichtung anderer B3S hinsichtlich SzA-Umsetzung: Dokumente explizit aus anderen KRITIS-Sektoren, deren Ausgestaltung jedoch gewisse Parallelen zur stationären Gesundheitsversorgung aufweist.

[12] https://www.enisa.europa.eu/topics/cybersecurity-policy/nis-directive-new/minimum-security-measures-for-operators-of-essentials-services

[13] https://informatik.th-brandenburg.de/forschung-und-kooperation/schwerpunkte-und-themen/medsec-it-sicherheit-in-der-medizinischen-versorgung/

Zur Analyse wurden ausschließlich verabschiedete Dokumente herangezogen. Punktuell wird innerhalb der nachfolgenden Unterabschnitte auf bedeutsame Entwicklungen oder Aktualisierungen hingewiesen.

Die Auswertung ist gesamtheitlich über den vollständigen Prozess zur Erarbeitung von Handlungsempfehlungen in Abschnitt 4.3 dokumentiert.

4.2.1 Standards mit Bezug zu SzA (sektorübergreifend)

Im Allgemeinen sind schon recht weitreichend Sicherheitsanforderungen zur Umsetzung von Angriffserkennung in einschlägigen Standards und sonstigen Handreichungen abgebildet. Die Analyse auf dieses Dokument fokussiert sich primär auf die Sachverhalte, die für SzA außerhalb von klassischen IT-Systemen herangezogen werden sollten.

Nicht aufgeführt ist an dieser Stelle der IT-Grundschutz des BSI (BSI 200-1 /-2 und -3 sowie insbesondere das Kompendium), weil dieser bereits integral in der Orientierungshilfe SzA mit aufgegriffen ist.

Die ISO 27002 (Informationssicherheit, Cybersicherheit und Schutz der Privatsphäre – Informationssicherheitsmaßnahmen) beschreibt allgemein und umfassend die Ausgestaltung eines ISMS und liefert daher naturgemäß wenig Detail zur Ausgestaltung einzelner Anforderungen wie bspw. im Zusammenhang mit SzA. Allerdings lassen sich den Abschnitten 8.15. und 8.16 Übersichten zur Auswahl von relevanten Inhalten für SIEM und IDS entnehmen, wenngleich diese sich nicht auf die Gesundheitsversorgung fokussieren.

Die Standards der Reihe ISO 27033 beschreiben konkret die sichere Ausgestaltung von IT-Netzwerken. In 27033-1 [23] wird in Abschnitt 8.5 vorgeschlagen, welche Ereignisse von Netzgeräten herangezogen und ausgewertet werden sollten. Ebenfalls in Abschnitt 8.5 allerdings von 27033-2 [24] wird dies noch einmal konkreter aufgegriffen und *syslogs, SNMP information, IDS/IPS alerts,* und *flow information* als relevante Informationen für security information monitoring herausgestellt.

Die Reihe der ISO 27035 (Information security incident management) beschreibt umfassend und ausführlich, wie der Umgang mit sicherheitsrelevanten Ereignissen und Sicherheitsvorfällen gestaltet und gesteuert werden kann. „Part 1: Principles and process" [25] ordnet das Thema ein und gibt einen soliden Überblick über Phasen und Rollen beim Information security incident management (SIM). Das SIM wird in die 5 Phasen *Plan and prepare; Detection and reporting; Assessment and decision; Responses; sowie Lessons learnt* aufgeteilt. In „Part 2: Guidelines to plan and prepare for incident response" [26] werden konkretere Handreichungen zu den Phasen eins und fünf bereitgestellt; so finden sich in Abschnitt 7 Rollenvorschläge für die Besetzung von Monitoring- und Reaktionsteams wieder. In den Anhängen B und C werden Beispiele sicherheitsrelevanter Ereignisse in Kategorien und Templates für Formulare mit thematischem Bezug vorgehalten. „Part 3: Guidelines for ICT incident response operations" [27]

behandelt die verbleibenden drei Phasen und schlüsselt diese weiter auf. In den Unterabschnitten 7.3.1 sowie 7.3.2 finden sich Hinweise zum konkreten Umgang mit Threat Intelligence sowie Verweise auf entsprechende Quellen wieder. Abschnitt 10.5 schlägt Tools für die unterschiedlichen Analysebereiche des SIM vor; Abschnitt 12.4 hält Formate und potentielle Quellen für „Information Sharing" bereit.

Aus der Standardreihe IEC 62443 („Industrielle Kommunikationsnetze – IT-Sicherheit für Netze und Systeme", vgl. Abbildung 8) wurden zwei Standards hinsichtlich relevanten Inputs für SzA geprüft. In der Normenreihe kommt das Konzept von „Security-Levels" zum Einsatz, welches das „Maß des Vertrauens, dass das IACS von Sicherheitslücken frei ist und in der beabsichtigten Weise funktioniert" beschreibt. Die auf industrielle Automatisierung zugeschnittenen Standards legen einen Fokus auf das Schutzziel Verfügbarkeit: „IT-Sicherheitsmaßnahmen dürfen sich nicht nachteilig auf wesentliche Funktionen eines hoch verfügbaren IACS[14] auswirken, außer dies wird durch eine Risikoeinschätzung unterstützt." IEC 62443-3-3 [28] legt anhand der sieben grundlegenden Anforderungen aus den allgemeinen Grundlagen nach IEC 62443-1-1 detaillierte technische Systemanforderungen an IACS fest. In den Abschnitten 6.10 und 10.3 wird deutlich, dass für OT in der Industrie (Automatisierungsanlagen) bereits seit Jahren der Export von Ereignisdaten in einschlägigen Formaten zu automatisierter Verarbeitung durch bspw. SIEMs und die Bereitstellung einer API zum programmierten Zugriff auf die Ereignisprotokolle normiert sind. IEC 62443-4-2 [29] ist strukturgleich zur 62443-3-3 und enthält Zusammenstellungen abgeleiteter Anforderungen zur detaillierten Abbildung von Systemanforderungen auf Subsysteme und Komponenten – sie wendet sich somit primär an Produktlieferanten und Dienstleister für IACS. Bezüglich SzA sind dem Dokument kaum mehr Konkretisierungen zu entnehmen; in den Abschnitten 6.10 und 10.4 gibt es punktuell konkrete Anhaltspunkte zu Inhalten von Ereignisdaten und Platzierung von Monitoring-Geräten in Industrieanlagen.

[14] Industrial Automation and Control Systems

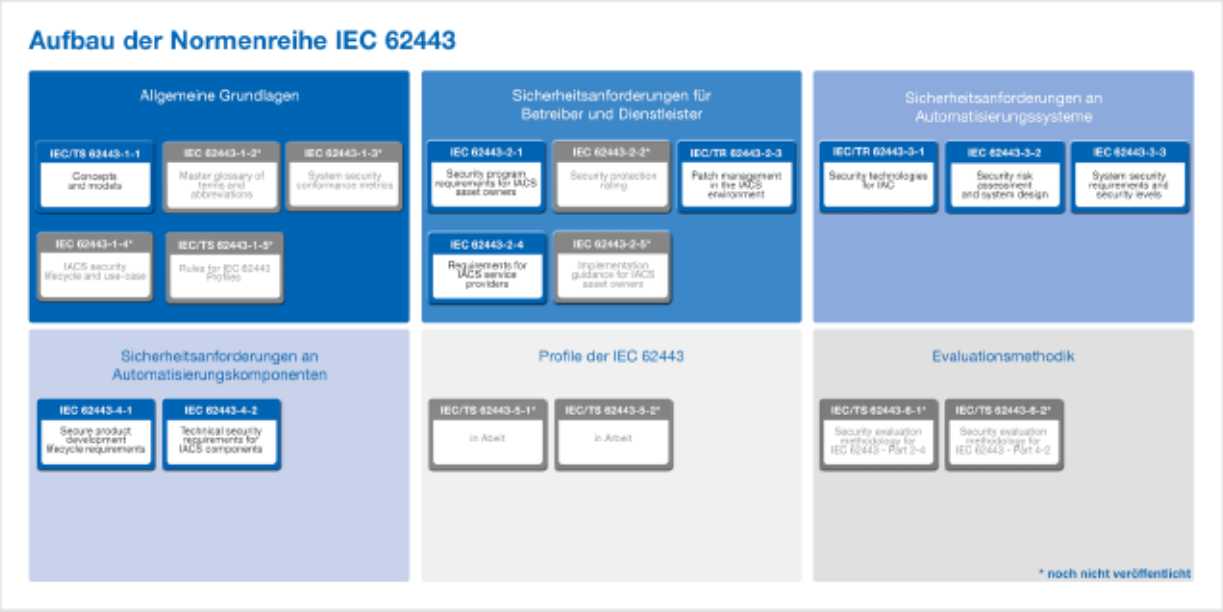

Abbildung 8: Aufbau der Normenreihe IEC 62443 (Quelle: DKE[15])

ISO 27039 (Selection, deployment and operations of intrusion detection and prevention systems (IDPS)) [30] erläutert im Detail die Einführung und den Betrieb von IDS. Da der Schwerpunkt der OH eher auf SIEM als auf IDS liegt, lassen sich schwerlich Konkretisierungen mit Relevanz für Krankenhäuser daraus ziehen. Allerdings werden hilfreiche Hintergründe eingeführt; in Abschnitt 5.5 beispielsweise wird das Zusammenspiel verschiedener Sicherheits- und Erkennungssysteme erläutert.

Das National Institute for Standards in Technology (NIST) hat bereits im Jahr 2007 die Special Publication (SP) 800-94 „Guide to Intrusion Detection and Prevention Systems (IDPS)" [31] veröffentlicht. Die Veröffentlichung erklärt die Eckpfeiler von Angriffserkennung, diskutiert die Automatisierung von Überwachung und Analyse von Ereignissen in Computersystemen und Netzwerken und beleuchtet die Notwendigkeit von IDS in der Sicherheitsinfrastruktur von Organisationen. Darüber hinaus deckt es insbesondere die Aspekte Systemdesign, Implementierung, Konfiguration, Sicherheit, Überwachung und Wartung ab. Das Dokument bietet detaillierte und praxisnahe Anleitungen für host-basierte, netzwerk-basierte und verhaltensbasierte Angriffserkennung. Außerdem gibt es eine Übersicht über weiterführende Detektionstechnologien (z. B. SIEMs). Im Jahr 2012 wurde eine Überarbeitung der SP 800-94 (Revision 1) eingeleitet, jedoch ohne eine endgültige Veröffentlichung eingestellt. Ein grundlegendes Update wurde von NIST Mitte 2022 für die Zukunft angekündigt, ist aber zum Zeitpunkt dieser Arbeit noch nicht veröffentlicht worden[16].

[15] https://www.dke.de/de/arbeitsfelder/industry/iec-62443-cybersecurity-industrieautomatisierung
[16] https://csrc.nist.gov/pubs/sp/800/94/final

NIST SP 800-61 (Rev. 2) „Computer Security Incident Handling Guide" bietet umfassende Anleitungen zur Reaktion auf Computersicherheitsvorfälle, mit einem Schwerpunkt auf der Reaktion und Erholung von Vorfällen [32].
Das NIST Cybersecurity Framework (CSF) besteht aus Standards, Richtlinien und bewährten Verfahren zur Verwaltung von Cybersicherheitsrisiken. Es ist anpassbar und wurde mit dem Ziel entwickelt, auf alle Arten von Unternehmen anwendbar zu sein. Am 26.02.2024 wurde eine grundlegende Überarbeitung des Frameworks (Version 2.0) veröffentlicht. Es verweist für die Systeme zur Angriffserkennung vor allem auf oben aufgeführte Special Publications und die ISO-Normen [33].

Abschließend ist in Tabelle 2 die Gesamtheit der untersuchten Dokumente im thematischen Umfeld „Standards mit Bezug zu SzA (KI-Sektor-Übergreifend)" aufgelistet.

Tabelle 2: Standards mit Bezug zu SzA

Quelle	Titel	Kurztitel	Jahr /Version	Ref.
BSI	Mindeststandard des BSI zur Protokollierung und Detektion von Cyber-Angriffen	MSt PuD	2023 2.0	[34]
ISO/IEC	Informationssicherheit, Cybersicherheit und Schutz der Privatsphäre – Informationssicherheitsmaßnahmen	ISO/IEC 27002	2022	[35]
ISO/IEC	Network security – Part 1: Overview and concepts	ISO/IEC 27033-1	2015 2nd ed.	[23]
ISO/IEC	Network security – Part 2: Guidelines for the design and implementation of network security	ISO/IEC 27033-2	2012 1st ed.	[24]
ISO/IEC	Information security incident management – Part 1: Principles and process	ISO/IEC 27035-1	2023 2nd ed.	[25]
ISO/IEC	Information security incident management – Part 2: Guidelines to plan and prepare for incident response	ISO/IEC 27035-2	2023 2nd ed.	[26]
ISO/IEC	Information security incident management – Part 3: Guidelines for ICT incident response operations	ISO/IEC 27035-3	2020 1st ed.	[27]
ISO/IEC	Selection, deployment and operations of intrusion detection and prevention systems (IDPS)	ISO/IEC 27039	2016 1st ed.	[30]
IEC	Industrial communication networks – Network and system security – Part 3-3: System security requirements and security levels (insb. Abschnitte 6.10 – 6.13, 7.11, 10.3 – 10.4)	IEC 62443-3-3	2019 (de)	[28]
IEC	Security for industrial automation and control systems – Part 4-2: Technical security requirements for IACS components (insb. Abschnitte 6.10 – 6.13, 7.11, 10.3 – 10.4)	IEC 62443-4-2	2019 1.0	[29]
NIST	Guide to Intrusion Detection and Prevention Systems (IDPS)	NIST SP 800-94	2007	[31]
NIST	Computer Security Incident Handling Guide (insb. Abschnitt 3.2 "Detection and Analysis")	NIST SP 800-61	2012 Rev. 2	[32]

Quelle	Titel	Kurztitel	Jahr /Version	Ref.
NIST	NIST Cybersecurity Framework 2.0	CSF 2.0	2024 2.0	[36]

4.2.2 Good Practice zu SzA in der stationären Gesundheitsversorgung

Konkret für die Gesundheitsversorgung wurden vom Deutschen Institut für Normung (DIN), von der ISO, von IHE International und vom US-amerikanischen Regelungsgeber Dokumente mit Bezug zu IT-Sicherheit herausgegeben.

Die DIN 80001 (Risikomanagement für IT-Netzwerke, die Medizinprodukte beinhalten) [37] beschreibt ein Risikomanagement für IT-Netze mit Medizintechnik, hält jedoch keine Konkretisierungen bzgl. SzA bereit. Mithin fokussiert das Dokument sehr auf den Prozess und seine Regelkreisläufe selbst. Zur DIN 80001 existieren neun sogenannte Ergänzungsnormen („Technical Reports"); für das Thema SzA wurde aus dieser Auswahl noch die „IEC/TR 80001 Part-2-2: 2012 – Guidance for the communication of medical device security needs, risks and controls" [38][17] mit geprüft, welche unter Abschnitt 5.2 das *Security Capability* „Audit controls – AUDT" führt. Dieses Dokument wiederum verweist für Details auf das „Audit Trail and Node Authentication (ATNA) Profile" [39] aus dem „IHE IT Infrastructure (ITI) Technical Framework, Volume 1"[18]. Das ATNA-Profil in Kombination mit der Spezifikation „Record Audit Event" aus dem Volume 2 des besagten IHE-Frameworks [40] liefert zahlreiche, detaillierte Anhaltspunkte zu Umsetzung und Inhalten von Systemprotokollierung in medizinischen Umgebungen.

Die ISO 27799 [41] beschreibt, wie ein ISMS nach ISO 27002 im Gesundheitswesen konkret umgesetzt werden kann. Dafür werden diverse Anforderungen aus der ISO 27002 hinsichtlich der Anwendung im Gesundheitsweisen konkretisiert. Gerade bei den Anforderungen im Zusammenhang mit SzA sind begrenzt Konkretisierungen erfolgt; jene Konkretisierungen beziehen sich wiederum vorrangig auf den Schutz der Daten in Gesundheitsinformationssystemen und den sicheren und nachvollziehbaren Zugriff auf die erhobenen Auditdaten selbst.

Insbesondere in den USA rückt die Angriffserkennung im Gesundheitsbereich verstärkt in den Fokus:
Der Health Insurance Portability and Accountability Act (HIPAA)[19] legt in seiner Sicherheitsregel Standards zum Schutz von Electronic Protected Health Information (ePHI) fest, die von betroffenen Einrichtungen und Geschäftspartnern eingehalten werden müssen. NIST hat Leitfäden und Werkzeuge entwickelt, um Gesundheitsorganisationen bei der Ausrichtung ihrer Cybersicherheitspraktiken auf das

[17] Lt. https://www.din.de/de/wdc-proj:din21:360211530 soll es in IEC/CD TS 81001-2-2 überführt werden
[18] https://profiles.ihe.net/ITI/TF/Volume1/index.html
[19] https://www.cdc.gov/phlp/publications/topic/hipaa.html

CSF zu unterstützen, die HIPAA-Anforderungen zu erfüllen. NIST SP 800-66 „Implementing the Health Insurance Portability and Accountability Act (HIPAA) Security Rule: A Cybersecurity Resource Guide" [42], deren zweite Revision im Februar 2024 veröffentlicht wurde, bietet betroffenen Einrichtungen Anwendungs- und Verständnishilfen für die Umsetzung der HIPAA Sicherheitsregel.

Der Cybersecurity Act forderte in Section 405(d) das US-Gesundheitsministerium (HHS) auf, gemeinsam mit der Gesundheitsbranche praktische Guidelines für Cybersicherheit zu entwickeln. Gemeinsam mit dem Healthcare and Public Health Sector Coordinating Councils (HSCC) hat es daraufhin den „Health Care and Public Health Sector Cybersecurity Framework Implementation Guide" [43] verfasst. Es handelt sich dabei um einen Leitfaden, der spezifische Schritte zur Implementierung des NIST Cybersecurity Frameworks im Gesundheitswesen bietet. Er umfasst verschiedene Anhänge und Tabellen, die unterschiedliche Aspekte der Implementierung abdecken. Beide Guidelines erleichtern die Anwendung des CSF im Kontext der HIPAA-Sicherheitsregel, SzA werden allerdings nicht in der Tiefe behandelt.

Im Rahmen des 405(d)-Programms werden derzeit gemeinsam mit der Gesundheitsbranche praktische Guidelines für Cybersicherheit entwickelt. Der aktuelle Projektstand kann auf der extra für dieses Programm erstellten Webseite[20] eingesehen werden.

In diesem Zusammenhang wurden 2023 die „Health Industry Cybersecurity Practices" [44] aktualisiert, die in ihren Technical Volumes aufgeteilt nach Größe der Betreiberorganisationen konkrete Hilfestellungen für die technische Umsetzung von IT-Sicherheitsmaßnahmen geben. Insbesondere im HICP Technical Volume 2 [45] für mittlere und große Organisationen sind zahlreiche technische Erklärungen hinsichtlich SzA im Gesundheitsbereich zu finden.

Abbildung 9: Health Industry Cybersecurity Practices (Quelle: HSCC [40])

Im Überblicksdokument [44] werden fünf Bedrohungen dargestellt, wovon eine RansomWare Angriffe und eine weitere Angriffe gegen vernetzte medizinische Geräte mit Auswirkungen auf die Patientensicherheit sind. In den technischen Volumes werden sodann entlang von zehn Cybersicherheitspraktiken („Cybersecurity Practices", CSPs) Gegenmaßnahmen zu diesen Bedrohungen beschrieben – für die Auswahl relevanter Maßnahmen wird nach Betreibergröße unterschieden (klein, mittel und groß), wobei die mittleren Unternehmen auch immer die Anforderungen an die großen Unternehmen mit prüfen sollten. Unter den Praktiken sind in Bezug auf SzA CSP 7 (Vulnerability Management) und CSP 8 (Security Operation Centers and Incident Response) relevant:

[20] https://405d.hhs.gov

- CSP 7. Vulnerability Management: Für kleine Unternehmen werden insgesamt vier Maßnahmen zu Schwachstellenscans, Umgang mit deren Ergebnissen und Patching regelmäßig und auf Hinweis der Hersteller an die Hand gegeben. Für mittlere Unternehmen werden diese beiden Themen weiter ausdifferenziert und es kommen Themen wie Klassifizierung von Systemen und Schwachstellen oder auch Konfigurations- und Änderungsmanagement hinzu. Großen Unternehmen werden noch weiterführende Anforderungen wie „Pentrationstesting" und Angriffssimulationen auferlegt.
- CSP 8. Security Operation Centers and Incident Response: Für kleine Unternehmen beschränkt sich dieser Punkt auf die beiden Anforderungen „Incident Response" (8.S.A) und „Information Sharing" (8.S.B), wobei Schwerpunkte der Notfallplanung für Cyberangriffe inklusiver entsprechender Rollen und der Vernetzung mit Organisationen oder Zentren für Informationsaustausch (ISAOs und ISACs) liegen. Bereits für Organisationen mittlere Größe wird explizit ein Security Operations Center (SOC, 8.M.A) (als Organisationseinheit, deren Mitglieder ihre Arbeitszeit vollständig der Cybersicherheit widmen) gefordert, für welches die Funktionen *Betrieb*, *Threat Intelligence* und *Incident Response* weiter detailliert sind. Die Anforderung *Incident Response* (8.M.B) legt deutlich detaillierter die notwendigen Vorbereitungen für eine Reaktion dar, wobei der Bereithaltung von Playbooks / Runbooks eine große Bedeutung beigemessen wird (Anhang G mit weiterführenden Ressourcen hierzu). Für große Unternehmen wird eine höhere Ausbaustufe eines SOC gefordert (Advanced Security Operations Center, 8.L.A), was sich beispielsweise in einem 24/7-Betrieb niederschlägt – in diesem Zusammenhang werden auch verschiedene Sourcing-Modelle für den SOC-Betrieb beleuchtet. Zusätzlich werden weitere Erkennungs- und Responsetechnologien wie User Behavior Analytics (UBA) oder Incident Response Orchestration in dieser Kategorie aufgeführt.

Dem Thema der Absicherung von netzwerkfähiger Medizintechnik wurde eine eigene Praktik (CSP 9) gewidmet, in dem die Möglichkeiten der Überwachung, ihre Grenzen und alternative Maßnahmen adressiert werden.

Für Health-Sector-spezifisches CTI-Sharing wurde das Health Sector Cybersecurity Coordination Center (HC3)[21] gegründet. Es befindet sich noch im Aufbau, veröffentlicht aber bereits regelmäßig akute und monatliche Bedrohungsmeldungen und gibt Sektoralarme heraus.

Insgesamt ist zu sagen, dass die US Behörden derzeit die über viele Veröffentlichungen verteilten Informationen hinsichtlich SzA konsolidieren und gleichzeitig eine große Initiative gestartet haben, konkrete, auf die Gesundheitsbranche angepasste Guidelines zu erstellen.

[21] www.hhs.gov/hc3

Abschließend ist in Tabelle 3 die Gesamtheit der untersuchten Dokumente im thematischen Umfeld „Good Practice zu SzA in der stationären Gesundheitsversorgung" aufgelistet.

Tabelle 3: Good Practice zu SzA in der stationären Gesundheitsversorgung

Quelle	Titel	Kurztitel	Jahr /Version	Ref.
DIN	Risikomanagement für IT-Netzwerke, die Medizinprodukte beinhalten	DIN EN 80001-1	2011	[37]
IEC	Guidance for the communication of medical device security needs, risks and controls	IEC/TR 80001 Part-2-2	2012	[38]
IHE	Audit Trail and Node Authentication (ATNA) Profile	ATNA Profile	2023	[39]
IHE	Record Audit Event [ITI-20]	ITI-20	2023	[40]
ISO	Informationssicherheitsmanagement im Gesundheitswesen bei Verwendung der ISO/IEC 27002	EN ISO 27799	2016	[41]
NIST	Implementing the Health Insurance Portability and Accountability Act (HIPAA) Security Rule: A Cybersecurity Resource Guide	NIST SP 800-66 Rev. 2	2024	[42]
HHS/HSCC	Health Care and Public Health Sector Cybersecurity Framework Implementation Guide		2023	[43]
HSCC	Health Industry Cybersecurity Practices: Managing Threats and Protecting Patients	HICP 2023	2023	[44]
HSCC	Technical Volume 2: Cybersecurity Practices for Medium and Large Healthcare Organizations	HICP Technical Volume 2	2023	[45]
HSCC	Technical Volume 1: Cybersecurity Practices for Small Healthcare Organizations	HICP Technical Volume 1	2023	[46]

4.2.3 Sichtung anderer B3S hinsichtlich SzA-Umsetzung

Im begrenzten Umfang wurden ebenfalls B3S aus anderen Branchen / Sektoren der Kritischen Infrastrukturen in Deutschland analysiert. Maßgeblich für die Auswahl war, dass innerhalb des vom B3S erfassten Geltungsbereichs ein hoher Anteil an Betriebstechnik (operational technology, OT) vorliegt.

Der B3S für die Verteilung von Fernwärme [47] ist auf die ISO27001 ausgerichtet und nimmt die Betreiber hinsichtlich der konkreten Ausgestaltung von Anforderungen in die Pflicht. Zum Thema SzA existiert unter 4.3 ein eigener Abschnitt mit dem Titel „Vorfallerkennung und -bearbeitung". Unter Abschnitt 4.7 „Technische Informationssicherheit (Maßnahmenkategorien)" werden Themen wie IDS/ IPS, SIEM und SOC auch noch einmal konkreter aufgegriffen.

Der B3S für Wasser / Abwasser [48] stellt im Gegensatz dazu vollständig auf den IT-Grundschutz ab. Bezüglich der Anforderungen erfolgt ausschließlich eine Referenzierung auf Anforderungen aus dem IT-Grundschutz-Kompendium. Die ca. 20 Anforderungen konkret zum Thema SzA sind in der Kategorie

„Angriffserkennung und Reaktion" zusammengefasst. Umsetzungshinweise werden zu Teilen direkt aus dem Kompendium referenziert, teilweise werden diese auch auf die Branche angepasst. Weitere Umsetzungshinweise werden von der Branche erstellt.

Abschließend ist in Tabelle 4 die Gesamtheit der untersuchten Dokumente im thematischen Umfeld „Sichtung andere B3S hinsichtlich SzA-Umsetzung" aufgelistet.

Tabelle 4: Branchenspezifische Sicherheitsstandards außerhalb der Gesundheitsversorgung

Branche / Sektor	Titel	Kurztitel	Jahr /Version	Ref.
Fernwärme	Branchenspezifischen Sicherheitsstandard für die Verteilung von Fernwärme	B3S VvFw	2021 1.1	[47]
Wasser / Abwasser	IT-Sicherheit – Branchenstandard Wasser/Abwasser	B3S WA 2023[22]	2023	[48]

4.3 Ergebnisse

In Tabelle 5 sind die Ergebnisse aus allen vorhergehenden Analyseschritten konsolidiert. Die Tabelle ist mit ihren vier inhaltlichen Bereichen (übergreifende Ergebnisse, Protokollierung, Detektion, Reaktion) entlang der Struktur der OH SzA strukturiert. In Kapitel 4.4 werden anschließend die abgeleiteten Handlungsempfehlungen in einer eigenen Tabelle dokumentiert.

Tabelle 5: Ergebnisse aus Befragung, Expertengesprächen und Dokumentenauswertung

Nr.	Ergebnis	Erläuterung / Konkretisierung
Übergreifende Ergebnisse		
1.	Die Umsetzung von SzA in Krankenhäusern bringt besondere branchenspezifische Herausforderungen mit sich. Diese sind jedoch nicht primär technischer Natur; vielmehr sind die Umsetzenden durch das regulatorische und ökonomische Umfeld herausgefordert.	Als Herausforderungen bei der Umsetzung von IT-Sicherheit im Allgemeinen aber auch SzA im Konkreten wurde wiederholt der Konsolidierungsdruck bei Krankenhausstandorten angeführt. Zudem führten die Rahmenbedingungen zu niedrigen Geschwindigkeiten bei Investitionsentscheidungen und deren Umsetzungen. Nicht zuletzt fehlt qualifiziertes Personal für IT-Sicherheit. Das Projektteam beobachtete zudem, dass Gegebenheiten der Krankenhäuser in Kombination mit Durchsetzung gesetzlicher Pflichten seitens des BSI ggü. den Betreibern das Zusammenwirken zwischen Betreibern und BSI stören und zu Vertrauensverlusten führen.
2.	Konkrete – insb. technische – Standards und Handreichungen zur Umsetzung von IT-Sicherheit im Allgemeinen und SzA im Konkreten in Krankenhäusern existieren	Mit dem B3S MV wurde regelrecht Pionierarbeit geleistet, deren Ergebnis durchweg auf hohe Anerkennung stößt. Der Arbeitsmodus aus der Vergangenheit zur Erarbeitung des B3S MV (alle

[22] Entwurf, welcher noch nicht vom BSI bzgl. Eignung geprüft wurde und insofern u.U. noch Änderungen unterliegt.

Nr.	Ergebnis	Erläuterung / Konkretisierung
	bislang kaum und befinden sich – international – erst in Erarbeitung.	beteiligten Akteure zusammenbringen und Expertenmeinungen gegeneinander abwägen) wird weithin als erfolgreich angesehen. Diverse existierende Standards können in deutschen Krankenhäusern nicht oder nur bedingt angewendet werden. ISO27799 wurde von der Branche als unpassend verworfen, da insb. zentralisierte Strukturen auf nationaler Ebene vorausgesetzt werden. DIN 80001 fokussiert Safety (statt Security).DIN 13080 („Gliederung des Krankenhauses in Funktionsbereiche und Funktionsstellen") [49], lieferte notwendigen Input für Funktionsstellen / Funktionsdienste im Krankenhaus. Die US-Behörden machen derzeit große Fortschritte bei der Adaption der IT-Sicherheits-Frameworks und –Guidelines auf den Gesundheitsbereich. Aktuelle Veröffentlichungen sind zwar noch nicht komplett ausgereift, bieten aber schon jetzt deutlichen Mehrwert.
3.	Zur Perspektive von B3S bzgl. Stand der Technik IT-Sicherheit in KRITIS-Bereichen herrscht auf Seiten der Betreiber Unklarheit und teilweise Resignation.	Das BSI sieht enormes Potential in B3S und will das Thema zukünftig noch weiter stärken. Das BSI bietet in diesem Zusammenhang auch an, Erkenntnisse aus dem Mängelmonitoring (bspw. Häufungen von Mängeln) zur Fortschreibung zu überlassen. Teile der Branche zweifeln an der formalen Wirkkraft des B3S insbesondere mit Blick auf Nachweispflichten und die verbindliche Vollständigkeit der Anforderungen. Eine saubere Kommunikation und Verwendung der Begrifflichkeiten ist hier zuträglich: Der in der Branche entwickelte Prüfnachweisplaner (Excel-Tool) beispielsweise kann durchaus hilfreich sein, ist aber schon formal nicht Bestandteil der B3S-Eignungsprüfung durch das BSI.
4.	Für die Anforderungen zur Ausgestaltung von SzA gemäß IT-Sicherheitsgesetz stehen kurzfristig keine substanziellen Änderungen an.	Die Anforderungen zu SzA werden von der expliziten OH SzA in die übergreifende OH B3S integriert werden; hierbei handelt es sich jedoch vielmehr um einen Formalismus – inhaltliche Anpassungen zum Thema stehen nicht zu erwarten.
5.	Ein B3S kann vom BSI auch dann als geeignet festgestellt werden, wenn eine Mapping-Tabelle fehlt und somit der Nachweis zur Abgeltung aller relevanten Anforderungen nicht beigefügt ist.	In der Konsequenz muss dieser Nachweis jedoch von jedem einzelnen Betreiber geleistet werden.
6.	Kleinere (nicht-KRITIS) Häuser sind durch das Thema SzA noch stärker herausgefordert als große Betreiberorganisationen; mithin fallen sie jedoch durch die Regelungen nach SGB V ebenfalls unter SdT-Pflichten zu IT-Sicherheit.	Die der deutschen KRITIS-Regulierung für IT-Sicherheit zu Grunde liegende Systematik mit 500.000 versorgten Personen (welche wiederum auf anlagenspezifische Schwellenwerte umgerechnet wurde) stößt bei Krankenhäusern an ihre Grenzen. Insofern ist der Vorstoß des BMG zur Ausweitung der SdT-Pflichten auf alle Krankenhäuser (SGB V nachvollziehbar. Methodisch ließe sich seine Abstufung von Regelungen entlang der KRITIS-Grenze in einem B3S umsetzen (bspw. SzA als MUSS für KRITIS und als SOLL für Nicht-KRITIS). Dies endet letztendlich in ethischen Fragen, ob in „kleineren" Häusern tatsächlich ein geringerer Schutz gegen Cyberbedrohungen gerechtfertigt ist.

Nr.	Ergebnis	Erläuterung / Konkretisierung
		In den USA wurden Dokumente mit abgestuften Regelungen zur IT-Sicherheit in Abhängigkeit von der Betreibergröße bereits veröffentlicht [46] [45]. MITRE bietet in Abschnitt 3.4 eine allgemeine Orientierung, welche Aufgaben IT-Security Teams ab welcher Größe in welcher Qualität erbringen können [50]. Sowohl die DKG als auch das Bayrische Landesamt für Sicherheit in der Informationstechnik[23] haben Arbeitshilfen und Dokumente herausgegeben, die sich zumindest teilweise primär an Betreiber unterhalb der KRITIS-Schwellenwerte richten.
7.	Ein Vordringen von Cyberangriffen in die Medizintechnik ist zwar nicht der Regelfall, lässt sich aber in Einzelfällen bereits heute beobachten.	Für die Zukunft steht hier eine Verschärfung zu erwarten, da sich, dem allgemeinen Trend folgend, Bedrohungen von der IT auf die OT ausweiten. Eine Analyse zur Cyber-Resilienz von Krankenhäusern des US-Gesundheitsministeriums[24] stellte 2023 zudem einen Trend in der Motivation von Angreifern fest. Neben finanziell motivierten Angreifern gab es solche, die einen Reputationsschaden verursachen oder das Vertrauen der Öffentlichkeit erschüttern wollten. Insbesondere für derart motivierte Angreifer dürften Medizingeräte besonders interessant sein.
8.	Gerade in kleineren Einrichtungen werden IT-Security Analysten mit der Auswertung von IT-Security Incidents aus den Erkennungssystemen nur für die eigene Organisation allein nicht ausgelastet sein.	Es muss bereits in der Planungsphase festgelegt werden, welche Aufgaben das IT-Security Team übernehmen soll und insbesondere, welche Aufgaben nicht. Aus den Aufgaben kann die notwendige Personalstärke abgeleitet werden. Eine Orientierung bietet das Capability-Template (Tabelle 4) von MITRE [50]. Die Auslastung der IT-Security Analysten muss zusätzlich mit den Betriebszeiten der Sicherheitsorganisation zusammen betrachtet werden.
9.	Die bisherigen Versionen des B3S für Krankenhäuser haben sich im Ergebnis des intensiven Diskussionsprozesses auf verschiedene Grundlagen (IT-Grundschutz und ISO27000 nativ) abgestützt.	Andere Branchen / Sektoren beziehen sich mit ihren B3S klar auf eine konkrete Vorgehensweise (so Wasser / Abwasser auf den IT-Grundschutz und Fernwärme auf die native ISO 27000). Ein B3S muss den Spagat zwischen ausreichend Umsetzungsspielraum für die Betreibenden und ausreichender Konkretheit der Anforderungen für das eigentliche Umsetzen leisten.
10.	Die technische Ausstattung der Krankenhäuser weist eine außerordentliche Heterogenität auf.	Auch andere KRITIS-Sektoren weisen Unterschiede in der Ausgestaltung ihrer System- und Netzlandschaften aus. Jedoch zeugt die Ausgestaltung der Systemlandschaften zwischen den Krankenhäusern – gerade in Bezug auf MT – von einer besonders hohen Heterogenität. Verschiedene Anbieter von Medizinprodukten für unterschiedliche Einsatzzwecke kommen zum Einsatz; der Grad an Standardisierung und Vereinheitlichung ist äußerst gering.

[23] https://www.lsi.bayern.de/kritis/krankenhaeuser/index.html
[24] https://405d.hhs.gov/Documents/405d-hospital-resiliency-analysis.pdf

Nr.	Ergebnis	Erläuterung / Konkretisierung
11.	Dienstleister (MSSP) bauen aktuell erst Branchenexpertise für Krankenhäuser auf.	Es zeichnet sich eine zeitliche Korrelation zwischen der gesetzlichen Anforderung nach SzA durch das zweite IT-Sicherheitsgesetz zu Mai 2023 und den Dienstleistungen, die im Bereich Managed Security Services für Krankenhäuser angeboten werden, ab. Historisch existieren einerseits IT-Systemhäuser, die sich auf die Gesundheitsversorgung spezialisierten, und Managed Security Service Provider (MSSP) andererseits, die Erfahrungen konkret mit dem Thema SzA aber keine Branchenkenntnis haben. Erst langsam bilden sich am Markt Akteure heraus, die konkret SzA in Krankenhäusern auch als Referenz ausweisen können.
12.	Es gibt in der Branche eine Kontroverse, ob auf Grund von §203 StGB Patientendatenverarbeitung im Kontext SzA nur in Deutschland stattfinden darf.	MSSPs agieren oftmals global; gerade auch um 24/7-Bearbeitung abzubilden. Eine solche Besonderheit der Gesundheitsbranche würde die Auswahl der MSSPs deutlich einschränken oder umfangreiches Pre-Processing der übermittelten Daten erfordern. Aus dem Wortlaut des Paragraphen[25] lässt sich eine Anforderung an die Lokation zur Datenverarbeitung nicht unmittelbar ableiten.
13.	Die Anwendung der MITRE-ATT&CK-Matrix ist noch sehr divers in der Branche; sie gewinnt aber zunehmend an Relevanz.	Sofern in Nutzung, dient die MITRE-ATT&CK-Matrix in vielen Fällen zur Priorisierung und Vollständigkeitsprüfung von Use Cases, die von den Erkennungssystemen berücksichtigt werden sollen.
14.	Die Gematik als Betreiberin der Gesundheits-TI in Deutschland trifft Vorbereitungen für IT-Angriffserkennung.	Für die Ausgestaltung der Vorbereitungen ist nicht ausgeschlossen, dass auch Systeme in den Krankenhäusern selbst mit einbezogen werden könnten.
Protokollierung		
15.	Für netzwerkfähige Medizintechnik lässt sich im Allgemeinen keine Systemprotokollierung aktivieren.	Eine überblicksartige Stichprobe an der TH Brandenburg unterfüttert die These, dass netzwerkfähige Medizintechnik selten in der Lage ist, auskömmlich Systemprotokollierung zu erzeugen und zu exportieren. In Einzelfällen wird ein solches Feature gegen zusätzliche Vergütung angeboten. Hinzu kommt, dass lange Lebenszyklen der eingesetzten Technik das Problem noch auf einige Jahrzehnte verlängern.
16.	Die Umsetzung der Netzsegmentierung in den Krankenhäusern weist eine sehr unterschiedliche Tiefe und auch sehr unterschiedliche Reifegrade auf.	Abseits der allgemeinen Erfordernisse zur Absicherung der Infrastrukturen durch Segmentierung der Netze erschwert diese Diversität auch ein einheitliches Vorgehen bei SzA.
17.	Für Medizintechnik fordern die Hersteller (/Wartungspartner) oftmals Fernzugriff.	Teilweise müsse als Anforderung der Hersteller oder Wartungspartner der Fernzugriff permanent aktiv sein. Hersteller von Medizintechnik fordern unterschiedliche Fernzugriffslösungen ein, sodass – selbst innerhalb von großen Häusern – keine Einheitlichkeit besteht.

[25] https://www.gesetze-im-internet.de/stgb/__203.html

Nr.	Ergebnis	Erläuterung / Konkretisierung
18.	Der Umgang mit Datenschutz – insb. hinsichtlich der Speicherfristen – ist äußerst divers.	Aus Sicherheitssicht besteht insbesondere zum Zweck der forensischen Analyse die Notwendigkeit, Ereignisdaten lange aufzubewahren (um beispielsweise nach Wochen oder gar Monaten, als ein Angriff detektiert wurde, die Spuren bis zur Erstinfektion zurückverfolgen zu können). Dies ist auszubalancieren insbesondere gegen Anforderungen aus dem Datenschutz und zu Teilen auch gegen Kosten für Speicher.
19.	Cloud-Dienste sollten zu deren Schutz in SzA mit eingebunden werden	Die Erhebung des Ist-Zustands hat ergeben, dass 32% der Antwortenden Cloudtechnologie einsetzen. Die aktuelle Bedrohungslandschaft zeigt, dass auch solche Systeme nicht frei von Schwachstellen und Fehlkonfigurationen sind und Ziel von Angriffen werden können. Wenngleich in der OH SzA die Absicherung von Cloudtechnologie nicht explizit beschrieben wird, sollte sie ebenfalls in die Protokollierung einbezogen werden. Dies gilt insbesondere, wenn sie Teil der kritischen Dienstleistung ist.
20.	Die Umfrage hat gezeigt, dass sehr unterschiedlich viele Quellen und sehr unterschiedlich große Log-Volumina in das SIEM eingespeist werden und das zu einer großen Divergenz in SRE und qSRE führt.	Grundsätzlich gibt es zwei valide Vorgehensweisen bei der Erschließung von Logquellen: „tune up from zero or tune down from everything". Beide Ansätze sind vertretbar und haben jeweils Vor- und Nachteile, die in MITRE Strategie 7.1.6 detailliert dargelegt werden [50]. Es sollte berücksichtigt werden, dass beide Ansätze die Gefahr bergen, dass sicherheitsrelevante Ereignisse von den IT-Security-Analysten zunächst nicht erkannt werden. Entweder, weil das betroffene System noch nicht (vollständig) erschlossen ist, oder, weil sich das IT-Security-Team noch nicht in der Fülle der Informationen zurechtfindet. Es ist daher essentiell, dass sich die IT-Security-Analysten sequentiell mit den angebundenen Systemen und deren SRE vertraut machen.[26]
21.	In den internationalen Standards gibt es zahlreiche Hinweise darauf, welche Ereignisse auf System- oder Netzebene für ein Security-Monitoring herangezogen werden sollten.	Einige der geprüften Unterlagen liefern konkreten Input dazu, welche Ereignisse potentiell als relevant anzusehen sind. Dazu zählen insb.: - ISO27002 [35] mit Vorschlägen für SIEM-relevante Ereignisse, Überwachungsmaßnahmen für Protokollanalyse auf Netzebene, sowie Inhalte für IDS - ISO27799 [41] mit wenigen konkretisierenden Protokollierungsinhalten für Gesundheitsinformations-Systeme - ISO/IEC27033-1 [23] für Protokollierungsinhalte auf Netzebene - ISO/IEC27035-3 [27] mit einem Überblick über Methoden und konkreten Erkennungstechnologien. - IEC 62443-4-2 [29] zu relevanten Ereignissen in industriellen Umgebungen. - MITRE SOC Strategien 7.3 zu SRE auf Hostebene und 7.4 auf Netzebene [50].

[26] Vergleiche [50] S. 179f

Nr.	Ergebnis	Erläuterung / Konkretisierung
		- Center for Internet Security mit sehr konkreten Umsetzungshinweisen bis auf Command-Ebene für eine Vielzahl von verbreiteten Systemen.[27]
Detektion		
22.	Eine technische Umsetzung von SzA ohne eine zentrale Auswertungsinstanz für Ereignisdaten wird ggü. dem BSI mit massivem Argumentationsaufwand bzgl. SdT verbunden sein.	Die OH-SzA legt fest, dass alle sicherheitsrelevanten Protokoll- und Protokollierungsdaten an für den jeweiligen Netzbereich zentralen Stellen gespeichert werden müssen. Dies wird vielfach mit einer SIEM-Pflicht ab der ersten Sekunde gleichgesetzt. Allerdings können insbesondere kleine IT-Security-Teams und SOCs im Aufbau zunächst auch mit einer Kombination von zentralem Log Management und EDR SzA umsetzen. MITRE weist darauf hin, dass ein Upgrade auf ein SIEM erst notwendig wird, wenn der Bedarf des IT-Security Teams die Möglichkeiten eines zentralen Logmanagement-Tools übersteigen. Insbesondere während der Aufbauphase können weitere Technologien eingesetzt werden, die ein IT-Security-Team unterstützen, Angriffe auf IT-Strukturen trotz noch nicht vollständiger Abdeckung zuverlässiger zu erkennen. Dazu gehören zum Beispiel Deception Systeme. Perspektivisch wird insbesondere ab einer bestimmten Größe eine Umsetzung von SzA ohne SIEM oder vergleichbare Technologie allerdings kaum zu rechtfertigen sein.
23.	Logdatenquellen von Endpoints und aus der Netzverkehrsanalyse haben unterschiedliche Schwerpunkte bei ihren Erkennungsfähigkeiten.	Bei der Auswahl der Logdatenquellen kann grundsätzlich angenommen werden, dass Daten eines Endpoints informativer sind als solche, die aus einer Netzverkehrsanalyse stammen – insbesondere, wenn es darum geht, einen Angriff zu bestätigen. Dagegen liegt der Vorteil von Netzverkehrsdaten darin, schnell feststellen zu können, ob und wohin sich Angreifer ausgebreitet haben und welche Systeme (zusätzlich) überprüft werden sollten.
24.	Über Schwachstellen in Medizinprodukten werden Betreiber regelmäßig über etablierte Kanäle informiert.	Hersteller oder Wartungspartner informieren über Schwachstellen in den entsprechenden Produkten. Weiterführende Informationen zu Bedrohungen und Angriffsmuster mit Fokus auf die Zielgruppe Krankenhäuser sind nicht etabliert.
25.	Es lässt sich eine enorme Spannbreite bzgl. der Qualität und Quantität von CTI in der Branche beobachten.	Es existieren kaum krankenhausspezifische Angebote; branchenübergreifende Quellen bieten aber durchaus branchenspezifische Einblicke: Die Enisa Threat Landscape (ETL) 2023 [51] beispielsweise gibt einen Überblick über Relevanz von RansomWare-Gruppen inklusive spezifischer Ausprägungen für Krankenhäuser. Gerade in den USA gibt es junge Aktivitäten zur Bereitstellung von auf die Gesundheitsversorgung zugeschnittenen CTI. In ihrer Bewertung zur Motivation von Angreifern (durchaus auch Sabotage) weichen diese von der Einschätzung der Bundesregierung bzw. des BSI („nicht erkennbar, dass die

[27] https://downloads.cisecurity.org/#/

Nr.	Ergebnis	Erläuterung / Konkretisierung
		Angreifer zwischen Krankenhäusern und anderen Wirtschaftsbeteiligten differenzieren, so dass eine Motivation zu einer Fokussierung auf Krankenhäuser nicht erkennbar ist" [52]) ab. Oftmals wird das Thema Threat Intelligence den Dienstleistern überlassen.
Reaktion		
26.	Eine vollständig automatisierte Reaktion auf Sicherheitsvorfälle lässt sich im Bereich der eigentlichen Leistungserbringung – zumindest in der aktuellen Ausgestaltung – nicht umsetzen.	Die Formulierung in der OH SzA schafft hierfür konkret die Möglichkeit für Ausnahmen: „In Netzen, wo die kritische Dienstleistung durch die Umsetzung nicht gefährdet wird, MUSS es möglich sein, automatisch in den Datenstrom einzugreifen, um einen möglichen Sicherheitsvorfall zu unterbinden." Die Reaktion auf einen Vorfall und deren Automatisierungspotenziale umfassen mehr als nur die Unterbrechung des Datenstroms; ISO27029 [30] hat dies systematisch aufgeschlüsselt und beispielsweise die Einsammlung von zusätzlichen Informationen explizit mit unter den Begriff „Active Response" gefasst.

4.4 Schlussfolgerungen und Vorschläge

Auf Basis der Ergebnisse aus Abschnitt 4.3 wurden Handlungsvorschläge abgeleitet und für jeden einzelnen Vorschlag – soweit jeweils möglich und sinnvoll – Formulierungsvorschläge für eine Überführung in den B3S erarbeitet. Hierfür wurden die Ergebnisse aus Abschnitt 4.3 vom Projektteam einzeln und im Zusammenhang systematisch durchgearbeitet. In Spalte „Ref. Ergebnisse" sind die Zeilen aus Tabelle 5 referenziert, welche für die Schlussfolgerung maßgeblich waren. Für die Spalte „Formulierungsvorschlag" erfolgte eine Feedbackrunde mit Vertretern der DKG und dem Branchensprecher des BAK MV (Workshop und schriftliche Kommentierung), da diese durch ihre Erfahrungen bei der Erstellung von bisherigen B3S-Versionen wertvolle Einordnung beisteuern konnten.

Nachfolgend sind in Tabelle 6 diese konkreten Schlussfolgerungen dokumentiert, wie mit den Ergebnissen aus Abschnitt 4.3 umgegangen werden sollte. Die Tabelle ist mit ihren vier inhaltlichen Bereichen entlang der Struktur der OH SzA strukturiert. In der Spalte „Formulierungsvorschlag" ist immer dann „n.a." eingetragen, wenn sich eine Schlussfolgerung nicht oder nicht sinnvoll mit einer Formulierung in einem zukünftigen B3S abbilden lässt.

Das Projektteam kann im Ergebnis der Analysen weder umfassend noch abschließend bestimmen, welche Verbindlichkeiten einzelne Anforderungen hinsichtlich ihrer Umsetzung haben müssen. Im Bericht ist dieser Umstand dadurch transparent gemacht, dass in der Spalte „Formulierungsvorschlag" innerhalb von Tabelle 6 die einschlägigen Modalverben („MÜSSEN", „SOLLTEN", „KÖNNEN") als Optionen aufgeführt sind. Eine abschließende Entscheidung muss sodann im Rahmen der Erstellung des B3S bei Würdigung dieser Vorschläge unter Hinzuziehung des BSI erfolgen.

Tabelle 6: Aus Ergebnissen abgeleitete Handlungsvorschläge

Nr.	Beschreibung der Schlussfolgerung	Ref. Ergebnisse	Formulierungsvorschlag
Übergreifende Vorschläge			
a.	Austausch auf politischer Ebene zw. Vertretung der umsetzenden Betreiberorganisationen und den zuständigen Ressorts (BMI und BMG), um für die umsetzenden Betreiberorganisationen einen Rahmen zu schaffen. Hierbei sollten die zuständigen Aufsichtsbehörden der Länder einbezogen oder zumindest berücksichtigt werden.	1	n.a.

Nr.	Beschreibung der Schlussfolgerung	Ref. Ergebnisse	Formulierungsvorschlag
	Die aktuelle, äußerst angespannte Bedrohungslage im Cyberraum insbesondere mit der Hauptbedrohung RansomWare, bei welcher Krankenhäuser im Branchenvergleich auch noch besonders betroffen sind, erfordern eine herausragende Priorisierung von Informations- und IT-Sicherheit. Insbesondere gilt es zu berücksichtigen, dass den aktuellen Bedrohungen im Cyberraum nicht nur mit einmaligen Investitionen sondern mit versiertem Personal begegnet werden muss: beim Beispiel SzA wird dies besonders deutlich, da die Erkennungstechnologien wie IDS und SIEM nur dann ihre Wirkung entfalten, wenn die Meldungen auch ausgewertet und die Systeme kontinuierlich kalibriert werden.		
b.	Sofern es die Doppelrolle des BSI (Regulierer und Kooperationspartner) zulässt, sollte auch für die Fortschreibung des B3S – und somit die tiefere Einbeziehung von SzA in diesen B3S – eine aktive Beteiligung seitens des BSI am Diskussionsprozess erfolgen. Die Informationen aus dem Mängelmonitoring des BSI auch zu SzA sollten bei der Fortschreibung des B3S gewürdigt werden.	1, 3, 4	n.a.
c.	Die internationalen Entwicklungen zu Handreichungen bzgl. IT-Sicherheit in Krankenhäusern (insb. in den USA) sollten eng beobachtet werden und in die Weiterentwicklung von B3S auch in Bezug auf SzA einfließen.	2	n.a.
d.	Dem fortgeschriebenen B3S sollte eine Mapping-Tabelle beigefügt werden, in welcher die Abgeltung der Maßnahmen aus der OH SzA dokumentiert ist. Von Ansprechpartnern aus Krankenhäusern wird oftmals der Wunsch geäußert, mit dem B3S „direkt anfangen zu können". Zusätzliche Nachweispflichten stünden dem entgegen.	5	n.a.
e.	Wenngleich Krankenhäuser unterhalb der KRITIS-Schwellenwerte nicht unter die Regelungen nach §8a IT-Sicherheitsgesetz fallen und somit auch die OH SzA keine unmittelbare Anwendung findet, sollte dort ebenfalls das Thema SzA mit berücksichtigt werden. Unter Abwägung der Risiken und Rahmenbedingungen muss ein Konsens herbeigeführt werden, in welchem Umfang die Anforderungen dort gelten sollen. Hierfür können auch die unter Vorschlag c ausgewiesenen internationalen Entwicklungen mit herangezogen werden.	6	n.a.

Nr.	Beschreibung der Schlussfolgerung	Ref. Ergebnisse	Formulierungsvorschlag
	Technisch und auch organisatorisch sind für Betreiber aller Größenordnungen Lösungen auch für SzA verfügbar.		
f.	Für den Betrieb derjenigen Krankenhäuser, die nicht zu den Kritischen Infrastrukturen gemäß BSIG gezählt werden, sollten Unterstützungsangebote (Handreichungen, Schulungsangebote) hinsichtlich der Einführung SzA bereitgestellt werden. Wie ausgeführt, sind diese bei dem Thema besonders herausgefordert – die Unterstützung des BSI fokussiert sich auftragsgemäß aber auf die Betreiber der Kritischen Infrastrukturen.	6	n.a.
g.	Die Betreiber sollten sich untereinander zu besonders relevanten branchenspezifischen Anforderungen an MSSPs austauschen.	6, 11	n.a.
h.	Bei der Entscheidung für den / die zugrunde gelegten Standard(s) zur Fortschreibung des B3S sollte der Internationalisierungsgrad der Adressaten berücksichtigt werden; bei einem maßgeblichen Anteil an international agierenden Akteuren sollte die ISO-Normenreihe stärkere Berücksichtigung finden.	9	n.a.
i.	Für die Fortschreibung des B3S sollten auch hinsichtlich der Ausgestaltung SzA Austausche mit Wissensträgern aus Autorenschaften anderer B3S durchgeführt werden.	9	n.a.
j.	Es sollte eine bewusste Sourcing-Entscheidung für das Thema SzA getroffen werden. Sofern ein Dienstleister einbezogen wird, sind klare Rollenzuordnungen und Schnittstellen erforderlich. Ohne eine Mitwirkung aus der Betreiberorganisation selbst wird sich SzA nicht umsetzen lassen. Auch gilt zu berücksichtigen, dass der Dienstleister mit fähigem Personal gesteuert und überwacht werden muss.	8, 11	Auch wenn die Umsetzung von SzA von einem Dienstleister unterstützt wird, MUSS/SOLLTE der Betreiber die Anforderungen auf Auftraggeberpflichten prüfen und entsprechend umsetzen. Hierbei MUSS/SOLLTE der Betreiber insbesondere bei Security Incidents (beispielsweise in der vertieften Fallbearbeitung) operativ mitwirken und den/die Dienstleister aktiv steuern.
k.	Bei der Beschaffung von Medizingeräten sollten MDS2-Dokumente verbindlich eingefordert werden, um die Transparenz bzgl. der Sicherheitseigenschaften zu erhöhen. Auch sollte die Software Bill of Materials (SBoM; sofern nicht bereits Teil von MDS2) angefordert und die genutzten Softwarekomponenten und Versionen in Hinsicht auf die Veröffentlichung von Schwachstellen überwacht werden.	15, 24	Bei der Beschaffung von Medizingeräten MUSS/SOLLTE eine Dokumentation über die Sicherheitseigenschaften (vorzugsweise in Form einer MDS2-Dokumentation) mit eingefordert werden. Ebenfalls SOLLTE die Software Bill of Materials (SBoM) in diesem Rahmen mit angefordert werden, wenn sie nicht Teil der MDS2 Dokumentation ist.

Nr.	Beschreibung der Schlussfolgerung	Ref. Ergebnisse	Formulierungsvorschlag
			Für Bestandsgeräte MÜSSEN/SOLLTEN die Betreiber die Bereitstellung von MDS2-Dokumenten prüfen. Für Medizingeräte MÜSSEN/SOLLTEN genutzte Softwarekomponenten und Versionen in Hinsicht auf die Veröffentlichung von Schwachstellen überwacht werden.
l.	Für IT-Sicherheitskomponenten wie SIEM (oder Varianten wie zentrales Logmanagement in Kombination mit EDR, XDR, …) sollte ein Eintrag ins Verzeichnis der Verarbeitungstätigkeiten vorgehalten werden. Hierfür könnte eine Vorlage für die Branche zur Verfügung gestellt werden.	18	Für das IT-Verfahren SIEM (oder Varianten wie EDR, XDR …) MUSS ein Eintrag ins Verzeichnis der Verarbeitungstätigkeiten vorgehalten werden, sofern darin personenbezogenen Daten verarbeitet werden.
m.	Zur Gematik sollte bezüglich SzA im Rahmen der Fortschreibung / Abstimmung des B3S Kontakt gehalten werden.	14	n.a.
Protokollierung			
n.	Die Betreiberorganisation muss konkret festlegen, in welcher Taktung und Reihenfolge Logdatenquellen an die zentrale Logdateninfrastruktur angebunden werden. Die Entscheidung sollte jedoch im Bewusstsein der Gefahren und jeweiligen Herausforderungen getroffen werden. Unerlässlich ist, dass sich das IT-Security-Team schrittweise mit den Quellen vertraut macht.	20	In einer zeitlichen Planung SOLLTE/KANN dargestellt werden, welche Protokollierungsdatenquellen zu welchen Zeitpunkten erschlossen sein sollen. Hierbei MÜSSEN/SOLLTEN die Integrationskapazitäten des IT-Security-Teams berücksichtigt werden.
o.	Es ist essentiell, die relevanten Logdatenquellen für die Detektion auszuwählen. Die Entscheidung hängt maßgeblich von der Organisation, ihrer Größe, ihren Systemen und (Cloud-)Anwendungen, sowie der Netzstruktur ab und lässt sich daher nicht pauschal vorwegnehmen. Eine Planung und ein Bewusstsein darüber, welche Angriffsszenarien für die Betreiberorganisation relevant sind, ist daher unerlässlich.	8, 10, 13, 16, 19, 23	Für die Priorisierung zur Auswahl der Protokollierungsdatenquellen MUSS/SOLLTE neben der Kritikalität der Systeme (für die kDL) die Bedrohungslage mit berücksichtigt werden.
p.	Da für IT gemeinhin Protokollierungs- und Detektionsmechanismen zur Verfügung stehen und bekanntgewordene Angriffe bislang hierauf einen Schwerpunkt setzten, sollte für die kDL relevante IT zuerst in die SzA-Ausgestaltung einbezogen werden. Insofern sollten über die	7, 20	Die für die kDL relevante IT MUSS/SOLLTE zuerst in die SzA eingebunden werden. Sodann SOLLTE umgehend die für die kDL relevante Medizintechnik und Versorgungstechnik eingebunden werden.

Nr.	Beschreibung der Schlussfolgerung	Ref. Ergebnisse	Formulierungsvorschlag
	Kritikalität von Systemen hinaus weitere Faktoren für die Einbeziehungsreihenfolge berücksichtigt werden.		
q.	Systeme und Anwendungen der kDL, die in die Cloud ausgelagert wurden, müssen ebenfalls überwacht werden. Bietet der Cloudprovider eigene Telemetrie- und Sicherheitsquellen an, können diese für die Überwachung genutzt werden.	19	Systeme und Anwendungen der kDL, die in die Cloud ausgelagert wurden, MÜSSEN/SOLLTEN ebenfalls in die Systeme zur Angriffserkennung integriert werden. Bietet der Cloudprovider eigene Telemetrie- und Sicherheitsquellen an, SOLLTEN/KÖNNEN diese für die Überwachung genutzt werden.
r.	Die Erhebung hat gezeigt, dass MT und VT hinsichtlich Sicherheitsereignisse fehlende, unzureichende oder nicht heranziehbare Protokollierungsfähigkeiten haben. Andererseits scheint dies nicht auf alle Geräte dieser Klassen zuzutreffen und zukünftig werden immer mehr Geräte dazu in der Lage sein. Daher ergibt sich für die Einbindung folgende Prüfreihenfolge: 1. Prüfen ob MT/VT protokollierierungsfähig ist und alle notwendigen Informationen in die zentrale Protokollierungsinfrastruktur einspeisen kann. 2. Falls nicht, prüfen, ob ein angebundenes Commodity Operating System (Windows/Linux; z. B. Steuerungsrechner) die notwendigen Informationen bereitstellen kann. 3. Falls nicht, müssen Ersatzmaßnahmen getroffen werden. Ein besonderer Fokus sollte hierbei auf der Protokollierung auf Netz-Ebene liegen. Die Netzebene sollte jedoch nicht nur als Fallback-Option gesehen werden.	15	Soweit möglich, MUSS/SOLLTE die Systemprotokollierung – auch bei MT & VT – aktiviert und mit sinnvollen Inhalten in die zentrale Auswertung einbezogen werden. Sofern die Technik eine Protokollierung auf Systemebene nicht oder nur unzureichend ermöglicht, MÜSSEN Ersatzmaßnahmen getroffen werden. Ersatzmaßnahmen sind beispielsweise die Einbeziehung eines an die MT angebundenen Commodity Operating Systems (z.B. Steuerungsrechner) in die SzA oder die Protokollierung auf Netzebene.
s.	Für die Auswahl konkreter, relevanter Ereignisse zur Erfassung in der zentralen Protokollierungsinfrastruktur wie beispielsweise SIEM sollten etablierte Good Practices herangezogen werden. Es ist wichtig zu betonen, dass jedes System und jede Anwendung eigene Logging-Capabilities mitbringt. Allgemeine Vorgaben können daher nur einen Rahmen bieten. Dieser kann aber eine wertvolle Entscheidungshilfe darstellen.	21	Die folgenden Ereigniskategorien SOLLTEN bei der Erfassung im SIEM berücksichtigt werden (Kategorien gemäß ISO27002 [35]): - a) erfolgreiche und abgelehnte Systemzugriffsversuche; - b) erfolgreiche und abgelehnte Versuche, auf Daten oder andere Ressourcen zuzugreifen; - c) Änderungen der Systemkonfiguration; - d) Nutzung von Privilegien;

Nr.	Beschreibung der Schlussfolgerung	Ref. Ergebnisse	Formulierungsvorschlag
			- e) Nutzung von Dienstprogrammen und Anwendungen; - f) Dateien, auf die zugegriffen wurde, und die Art des Zugriffs, einschließlich des Löschens wichtiger Dateien; - g) vom Zugriffskontrollsystem ausgelöste Alarme; - h) Aktivierung und Deaktivierung von Sicherheitssystemen wie Anti-Virussystemen und IDS; - i) Erstellung, Änderung oder Löschung von Identitäten; - j) Transaktionen, die von Benutzern in Anwendungen ausgeführt werden. In einigen Fällen handelt es sich bei den Anwendungen um einen Dienst oder ein Produkt, das von einem Dritten bereitgestellt oder betrieben wird. Für die Weitergabe von Protokollierungsinformationen von medizinspezifischen Netzgeräten an die zentrale Protokollierungsinstanz SOLLTEN/KÖNNEN mindestens die Ereignisse gemäß Absatz 3.20.4.1.1.1 des IHE Technical Frameworks, Volume 2 „Record Audit Event“[28] [40] einbezogen werden.
t.	Die Struktur der Netzsegmentierung sollte bei der Ausgestaltung SzA auf Netzebene berücksichtigt werden: bei ausreichend kleinen Netzsegmenten reicht unter Umständen eine Überwachung an den Netzübergängen.	16	Die Größe der Netzsegmente MUSS/SOLLTE bei der Umsetzung der Detektion insbesondere hinsichtlich der zusätzlichen Detektionssysteme berücksichtigt werden.
u.	Die Protokollierung auf Netzebene bietet insbesondere auf Grund fehlender Protokollierungsfähigkeit und großer Heterogenität Potentiale für die Branche, weil viele Geräte ausschließlich darüber und zudem mit vergleichsweise geringerem Aufwand überwacht werden können. Dabei unterscheidet die Analyse auf Netzebene viele verschiedene Technologien und Formate (u.a. NIDS, Netflow, Netzverkehrs-Metadaten, PCAP).	10	An Netzübergängen zwischen externen und internen Netzen MUSS eine Netzüberwachung stattfinden. An Netzübergängen zwischen internen Netzen sowie innerhalb von großen Netzen MUSS/SOLLTE geprüft werden, ob eine Netzüberwachung Mehrwert bietet, oder ob die Überwachung auf anderem Wege bereits ausreichende Abdeckung erreicht.

[28] https://profiles.ihe.net/ITI/TF/Volume2/ITI-20.html#3.20.4.1.1.1

Nr.	Beschreibung der Schlussfolgerung	Ref. Ergebnisse	Formulierungsvorschlag
	Jede dieser Optionen kommt jedoch mit Stärken und Schwächen. Das größte Risiko für die Branche liegen in einer tendenziell höheren False-Positive Rate als auf Host-Ebene, sowie einer eingeschränkten Fähigkeit derzeit verfügbarer Technologie, branchenspezifische Protokolle zu analysieren. Dennoch sollte die Netzüberwachung so ausgestaltet werden, dass Angreiferbewegungen identifiziert und nachverfolgt werden können.		
v.	Die Erhebung von Daten sollte derart ausgestaltet sein, dass die Erhebungsintensität im Bedarfsfall erhöht werden kann (bspw. durch zusätzliche Aktivierung von PCAP-Recording im konkret ausgewählten Bereich für den Ereignisfall).	15	Die Ausgestaltung der Protokollierung auf Netzebene MUSS/SOLLTE die datenschutzrechtlichen Rahmenbedingungen berücksichtigen. Für den Ereignisfall MUSS/SOLLTE vorbereitet sein, dass die Protokollierung erweitert (mehr Events, z. B. PCAP statt nur Verkehrsdaten) werden kann.
w.	Technologien zur Erkennung auf Netzebene (insb. NIDS und Netflow-Analyse) sollten zum Zwecke eines effektiven Einsatzes in Krankenhäusern auf die dort primär eingesetzten Protokolle (HL7, DICOM und weitere) erweitert werden. Diese branchenspezifische Adaption von Netzüberwachungs-Technologie sollte die Branche als Einheit nachfragen und einfordern. Der Bedarf ist absehbar und erste Anbieter mindestens auf dem Weg.	15	Bei der Beschaffung von SzA MÜSSEN/SOLLTEN Analysefähigkeiten der im Krankenhaus primär genutzten Protokolle (insb. HL7v2, HL7 FHIR und DICOM), Datenstrukturen (MIOs) und Technologien eingefordert werden.
x.	Zur Erfassung der in die Protokollierung einbezogenen Systeme sollten Automatisierungspotentiale (CMDB Discovery) geprüft und möglichst erschlossen werden. Entsprechende Systeme sollten hinsichtlich der Anwendbarkeit in Krankenhausumgebungen weiterentwickelt werden.	10	Für die Abbildung aller Geräte (inkl. MT und VT) mit ihrer Software und ihren Versionsständen in der Configuration Management Database (CMDB) KÖNNEN Automatisierungstools eingesetzt werden. Bei deren Beschaffung MUSS/SOLLTE darauf geachtet werden, dass medizinspezifische Produkte erkannt werden können.
y.	Es sollten Beschaffungsrichtlinien für verbindliche Verankerungen von Sicherheitsmerkmalen festgelegt und perspektivisch konsequent durchgesetzt werden. Im Konkreten sollten darin Anforderungen für SzA (z.B. Möglichkeit zum Log-Forwarding) abgebildet sein.	15	In den Beschaffungsrichtlinien MÜSSEN/SOLLTEN Sicherheitsanforderungen als MUSS-Anforderungen verankert werden. Im Konkreten MÜSSEN darin Anforderungen für SzA (z.B. Möglichkeit zum Log-Forwarding) explizit abgebildet sein.
z.	Fernzugriffe sollten aktiv durch die SzA überwacht und priorisiert ausgewertet werden. Dies gilt auch und insbesondere für die Hersteller / Wartungspartner, die auf die Medizingeräte zugreifen.	17	Fernzugriffe MÜSSEN/SOLLTEN aktiv durch die SzA protokolliert und priorisiert ausgewertet werden. Dies gilt auch und insbesondere für die Hersteller / Wartungspartner, die auf die Medizingeräte zugreifen.

Nr.	Beschreibung der Schlussfolgerung	Ref. Ergebnisse	Formulierungsvorschlag
aa.	Die von der SzA-Ausgestaltung verarbeiteten Daten sollten konkret hinsichtlich Datenschutz-Relevanz eingeschätzt werden. In diesem Zusammenhang ist ebenfalls zu prüfen, ob Patientendaten von der SzA-Ausgestaltung betroffen sind.	12	Die für SzA herangezogenen Protokollierungsdaten MÜSSEN/SOLLTEN aufgelistet werden. Für jeden Datensatz ist zu prüfen, ob a) Patientendaten und / oder b) personenbezogene Daten betroffen sind.
bb.	Zu sinnvollen Aufbewahrungszeiträumen für relevante Daten aus den SzA-Systemen sollte sich in der Branche ausgetauscht werden. Hierfür sollten neben Datenschutzvorgaben und internationalen Aufbewahrungsempfehlungen die bekannten Zeitläufe von Angriffen startend von der Erstinfektion bis zur Entdeckung (dwell time) herangezogen werden. MITRE empfiehlt (Tabelle 15) für Triage mindestens 2 - 4 Wochen und für Forensische Untersuchungen mindestens 6 Monate [50]. Das BSI hatte in seinem Mindeststandard zur Protokollierung und Detektion von Cyber-Angriffen in Version 1.0a [53] in Referenz auf das inkludierte Rahmendatenschutzkonzept (RDSK) die Speicherfrist aller Protokollierungsdaten auf 90 Tage festgelegt; sicherheitsrelevante Ereignisse könnten länger aufbewahrt werden. In der Folgeversion der Richtlinie von 2023 wurde dieser pauschale Ansatz jedoch aufgegeben; die Anforderung lautet nunmehr, dass „eine konforme Speicherfrist [...] identifiziert und angewendet werden muss" [34]. Der Bundesbeauftragte für den Datenschutz war 2020 in seiner Stellungnahme zum Entwurf des IT-Sicherheitsgesetz 2.0 ebenfalls auf das Thema Speicherfristen eingegangen und verwies diesbezüglich auf ein Grundsatzurteil des Bundesverfassungsgerichtes zum Thema Vorratsdatenspeicherung, wonach „eine Speicherungsdauer von sechs Monaten an der Obergrenze dessen ist, was unter Verhältnismäßigkeitserwägungen rechtfertigungsfähig ist" [54]. Weitere Indikatorik können Berichte von IT-Sicherheitsunternehmen liefern. Das Unternehmen Crowdstrike hatte für 2020 eine „dwell time" von durchschnittlich 79 Tagen angegeben [55]; Mandiant beobachtet jüngst eine Verkürzung der durchschnittlichen „dwell time" auf nur noch 10 Tage für das Jahr 2023 [56]. Krankenhäuser und insbesondere Betreiber von Kritischen Infrastrukturen sollten bei den Aufbewahrungszeiten aber ebenfalls berücksichtigen, dass sie sich gegen hochwertige Angriffe wie Advanced Persistent Threats (APT) schützen müssen – diese verweilen bekanntermaßen	18	Die Aufbewahrungsfristen für datenschutzrelevante Rohevents zur zentralen Speicherung MÜSSEN/SOLLTEN explizit festgelegt und eingehalten werden. Der Wert für die festgelegte Speicherfrist MUSS/SOLLTE einerseits Aufbewahrungsgrenzen aus dem Datenschutz berücksichtigen und andererseits die Erkennbarkeit und Analyse hochwertiger Angriffe wie Advanced Persistent Threats (APT) ermöglichen, da diese im Bereich der Kritischen Infrastrukturen berücksichtigt werden müssen.

Nr.	Beschreibung der Schlussfolgerung	Ref. Ergebnisse	Formulierungsvorschlag
	länger in IT-Infrastrukturen bis sie detektiert werden. Die Aufbewahrungszeiten sollten daher keinesfalls zu kurz gewählt werden. Zusätzlich sollten Anforderungen aus der gewählten Überwachungstechnologie in die Festlegung der Aufbewahrungszeiträume einbezogen werden. Insbesondere beim Einsatz von KI zur Überwachung kann die Funktionsweise der Software längere Aufbewahrungsfristen nötig machen.		
cc.	Zugriffe auf Systeme wie SIEM, DER und XDR sollten zur Nachvollziehbarkeit der Aktivitäten darin protokolliert und manipulationssicher aufbewahrt werden.	18	Zugriffe von Nutzenden auf SIEM, EDR, und XDR MÜSSEN/SOLLTEN protokolliert werden.
dd.	In einem Anonymisierungs- und Pseudonymisierungskonzept sollte dargelegt werden, wie die Personenbeziehbarkeit aufgelöst wird, um Anforderungen aus dem Datenschutz gerecht zu werden.	12	Anonymisierungs- und Pseudonymisierungskonzepte MÜSSEN/SOLLTEN um SzA erweitert werden, um darzulegen, wie die Personenbeziehbarkeit aufgelöst wird.
ee.	Die ersten Anhaltspunkte zur Dimensionierung der zentralen Protokollierungsinfrastruktur aus der Umfrage des Projektteams (ca. 0,5 – 17 GB pro eingebundenem System und Tag) sollten weiter gegen die Realität abgeglichen und zur Ausgestaltung von SzA herangezogen werden. Durch Pre-Processing und Data Feed Tuning kann der Umfang eingespeister Logdaten drastisch minimiert werden, ohne dass Informationen mit Sicherheitsrelevanz verloren gehen. Die Speicher- und Verarbeitungskapazität sollte skalierbar ausgestaltet sein und die Realität regelmäßig gegen die Planung angeglichen werden.	20	n.a.
ff.	Da die Topologie von Cloud Assets in der Regel verteilter ist, bietet es sich an, die Überwachung zum Endpunkt zu verschieben. Insbesondere bei nicht-IaaS Ressourcen sollten hierfür die Telemetriequellen und Sicherheitsmechanismen der Cloud-Anbieter verwendet bzw. in die Protokollierungsinfrastruktur eingebunden werden.	19	n.a.
Detektion			
gg.	Sofern der B3S in seinem Detaillierungsgrad fortgeschrieben wird, sollten die relevanten Technologien (wie bspw. ein SIEM) auch benannt werden. Dies sollte jedoch erst dann erfolgen, wenn die Technologien in der Branche auch wirklich erprobt sind.	22	Für SzA MÜSSEN zentrale Komponenten zur Auswertung von Ereignissen eingesetzt werden. Dafür SOLLTE ein SIEM (oder Varianten wie XDR) zum Einsatz kommen.

Nr.	Beschreibung der Schlussfolgerung	Ref. Ergebnisse	Formulierungsvorschlag
hh.	Der Betrieb einer zentralen Detektionsinfrastruktur erfordert über den initialen Aufwand hinaus langfristige Pflege. Bestehende Datenfeeds müssen überwacht, adjustiert und stabil gehalten werden. Queries müssen optimiert und neue Systeme und Anwendungen der Organisation angebunden werden. Nicht zuletzt muss das organisationsspezifische Wissen des IT-Security-Teams übertragen und erhalten werden. Dafür benötigt das IT-Security-Team dauerhaft Ressourcen.	1	Bei der Ressourcenausstattung für SzA MÜSSEN neben den initialen auch die regelmäßigen Bedarfe berücksichtigt werden.
ii.	Für den Übergang zu einer voll ausgereiften Detektionsinfrastruktur kann gut gepflegte und platzierte Deception-/Honeypot-Technologie unterstützend eingesetzt werden. Vorteile liegen besonders in der äußerst geringen False-Positive Rate und dem verhältnismäßig geringen Initialisierungs- und Pflegeaufwand.	22	Deception-/Honeypot-Technologie KANN unterstützend zur Angriffserkennung eingesetzt werden.
jj.	Zu den aktuellen Entwicklungen und Trends bei den Bedrohungen aber auch den Sicherheitslösungen sollten sich die Betreiber nicht nur regelmäßig informieren, sondern auch branchenweit untereinander austauschen, um entsprechende Weiterentwicklungen rechtzeitig im Rahmen von B3S-Fortschreibungen berücksichtigen zu können.	22	n.a.
kk.	In die Bewertung der Bedrohungslage für das betriebene Krankenhaus sollte aktiv Cyber Threat Intelligence (CTI) einbezogen werden. Hierbei sollten auch krankenhausspezifische internationale Entwicklungen wie beispielsweise die ISACs auf Ebene der EU[29] oder aus den USA[30] berücksichtigt werden.	25	Krankenhausspezifische Bedrohungsinformationen MÜSSEN/SOLLTEN erschlossen und in die Sicherheitsbewertung der Organisation einbezogen werden.
ll.	Betreiberorganisationen – unabhängig von der Größe – sollten sich ein eigenes Bild von der IT-Bedrohungslage machen und hierfür Informationen einsammeln und auswerten. Hierfür sollte auch der Austausch operationalisiert werden, in welchem regelmäßig ein Abgleich zur Bedrohungslage auf den drei relevanten Ebenen (operativ, taktisch und strategisch) erfolgt. In	13, 24, 25	Der Betreiber und Auftraggeber MUSS/SOLLTE sich selbst regelmäßig ein Bild von der Sicherheitslage machen. Über Berichte des MSSP hinaus SOLLTEN weitere Informationen eingesammelt und Austausche zwischen Betreibern operationalisiert werden.

[29] https://www.isacs.eu/european-isacs
[30] https://h-isac.org/europe/

Nr.	Beschreibung der Schlussfolgerung	Ref. Ergebnisse	Formulierungsvorschlag
	diesem Zusammenhang sollte sich auch zu Vorfällen von Krankenhäusern innerhalb von Deutschland vertrauensvoll ausgetauscht werden.		
Reaktion			
mm.	Effektive Reaktion beginnt mit der Planung, auf welche Ereignisse in welcher Art und Weise reagiert werden soll. Dafür ist das Wissen erforderlich, welche Ereignisse für die Betreiberorganisation relevant werden könnten. Dafür sollten Kategorien sicherheitsrelevanter Ereignisse sowie zugehörige Reaktionsmaßnahmen definiert werden. Auch sollten Melde- und Eskalationsketten etabliert werden. Dafür eignen sich Play- und Runbooks.	8	Zu Abläufen bei der Reaktion SOLLTEN/KÖNNEN Playbooks entwickelt und vorgehalten werden. Sofern diese zwischen Betreibern ausgetauscht werden, MÜSSEN diese zuvor um sensible Details bereinigt werden.
nn.	Die Reaktion auf Sicherheitsvorfälle und insbesondere deren Automatisierungsgrad müssen differenziert abgewogen werden. Sofern Schutzmaßnahmen im Bereich der eigentlichen Leistungserbringung im Krankenhaus zu Kollateralschäden führen können (wie beispielsweise relevante Eingriffe in den Datenstrom), sollte deren Autorisierung nicht vollautomatisiert erfolgen.	26	Die Reaktion MUSS/SOLLTE automatisiert werden, sofern keine Rückwirkung auf die kDL erfolgen kann.
oo.	Automatisierungspotentiale der Reaktion finden sich nicht ausschließlich bei der automatisierten Abwehr von Angreifern (z.B. durch Eingreifen in den Datenstrom). Vielmehr können Schritte und Abfragen automatisiert werden. Beispielsweise können bestimmte Alarme dazu führen, dass relevante Informationen (über vordefinierte Queries) automatisiert eingeholt und den Sicherheitsanalysten präsentiert werden (sog. „alert enrichment").	26	n.a.
pp.	Die Potentiale zur netztechnischen Entkopplung von Systemen untereinander und vom Internet sollten geprüft und dokumentiert werden. Sofern ein von einem Sicherheitsvorfall (z.B. Schadsoftwarebefall) betroffenes System nicht vom Netz getrennt werden kann (bspw. bildgebendes Verfahren in einer medizinischen Operation) sollte die Entkopplung eines möglichst kleinen Netzsegmentes zeitnah umgesetzt werden können, um einerseits den Krankenhausbetrieb nicht zu gefährden, andererseits die Ausbreitung der Schadsoftware und weitere Schäden (wie bspw. Ausleitung von vertraulichen Daten) zu unterbinden oder mindestens zu hemmen.	26	Für eine zeitnahe Reaktionsmöglichkeit MÜSSEN/SOLLTEN Kommunikationsbeziehungen analysiert und hinsichtlich ihres Netztrennungspotentials bewertet werden. Hierbei MÜSSEN/SOLLTEN für potentielle Betroffenheiten möglichst kleine Kommunikationsinseln vorgesehen werden.

Nr.	Beschreibung der Schlussfolgerung	Ref. Ergebnisse	Formulierungsvorschlag
qq.	Die Informationen zu Bedrohungen und Reaktionen könnten systematisch in Security Orchestration Automation and Response (SOAR)-Systemen gepflegt werden, um die Standardisierung zu unterstützen. Solche Systeme bieten aber erst in etablierten SOCs großen Mehrwert, insbesondere, wenn die Bedarfe die Möglichkeiten existierender SIEM oder Sicherheitsvorfall-Management Tools übersteigen.	26	Informationen und Reaktionen KÖNNEN in einem SOAR gepflegt werden, um auf eine Vereinheitlichung hinzuarbeiten, den Austausch zu erleichtern und – sofern relevant – eine Automatisierung technisch einfacher möglich zu machen.
rr.	Die Erhebung zu den Betriebszeiten des IT-Security-Teams ergab sehr starke Unterschiede zwischen den Betreibern. Grundsätzlich sollte der Betriebsmodus an die Größe und die Zielsetzung angepasst und bewusst getroffen werden. Zusätzlich kann es sinnvoll sein, Modelle zwischen 8x5 und 24/7 in Erwägung zu ziehen. Möglichkeiten sind beispielsweise eine verlängerte Präsenz des IT-Security-Teams an Wochentagen oder ein verstärkter Einsatz von MSSPs in Randzeiten.	8	Die Betriebszeiten zur Wahrnehmung von und Reaktion auf Sicherheitsereignisse MÜSSEN/SOLLTEN klar definiert und kommuniziert sein. Gerade für Zeiträume außerhalb der Hauptbetriebszeit des Betreibers KÖNNEN Dienste von MSSPs herangezogen werden.

5 Zusammenfassung und Ausblick

Das zweite IT-Sicherheitsgesetz gilt zum Abschluss dieses Projektes bereits seit ca. drei Jahren und die Betreiber Kritischer Infrastrukturen müssen Anforderungen an SzA bereits seit Mai 2023 umgesetzt haben. Auch wenn der aktuelle B3S MV in Version 1.2 bereits SzA-Anforderungen adressiert und die Betreiber schon spürbar an dem Thema gearbeitet hatten, kann das Projektergebnis einen wertvollen Beitrag zur weiteren Ausgestaltung von SzA in Krankenhäusern leisten. Nicht zuletzt ist die Umsetzung von SzA – wie Informationssicherheit im Allgemeinen – keine einmalige Angelegenheit, sondern ein ständiger Prozess, bei welchem sich die Ausgestaltung an die Rahmenbedingungen anpasst und gleichermaßen Reifegrade nach und nach erhöht werden.

Im Rahmen dieses Projekts wurde zunächst eine umfassende Erhebung des aktuellen Stands der SzA-Umsetzung bei Krankenhausbetreibern durchgeführt, um eine fundierte Basis zur Evaluierung zu schaffen. Die Ergebnisse dieser Befragungen wurden in einem Workshop mit dem BAK-MV Arbeitskreis B3S-Fortschreibung diskutiert, qualitativ eingeordnet und weiter ergänzt. Zur Vertiefung der Analyse wurden insgesamt 25 allgemeine und branchenspezifische Dokumente sowie Standards zu Best-Practices im Bereich SzA untersucht. Darüber hinaus wurden acht Experteninterviews durchgeführt, um branchenspezifische Besonderheiten und Herausforderungen bei der Umsetzung von SzA zu identifizieren. Aus den gewonnenen Erkenntnissen konnten 44 Schlussfolgerungen gezogen werden, wovon 30 zu konkreten Formulierungsvorschlägen entwickelt wurden.

Die Untersuchung zeigt, dass der Einsatz von SzA in Krankenhäusern grundsätzlich möglich ist. Während Betreiber bei der Umsetzung diversen Herausforderungen begegnen, sind diese häufig nicht technischer Natur. Durch die vorliegende, systematische Aufarbeitung des Themas SzA für Krankenhäuser in Deutschland, die branchenübergreifende Dokumente und Expertisen aus anderen KRITIS-Sektoren berücksichtigt, kann wertvoller Input auch über die Medizinische Versorgung hinaus geliefert werden. Einige Sachverhalte, wie die Einbeziehung von Cloud-Diensten in SzA, zeichnen sich im Rahmen der internationalen Good Practices bereits ab und wurden in die Empfehlungen aufgenommen, obwohl sie über die expliziten Anforderungen der OH SzA hinausgehen. Auch eine Skalierbarkeit der Sicherheitsanforderungen entlang der Betreibergröße ist bereits heute in Good Practices abgebildet. Die Bedenken und Sorgen der Branche wurden aufgenommen und detailliert beleuchtet, insbesondere die Möglichkeiten und Grenzen im Kontext von SzA bei Medizintechnik und Versorgungstechnik. Dies wurde auch mit der Regulierungsbehörde erörtert. Im Ergebnis wurde ein Vorgehensvorschlag entwickelt, wie die Einbeziehung von Medizin- und Versorgungstechnik trotz aller Herausforderungen gelingen kann.

Insgesamt steht die Diskussion über die Ausgestaltung von Systemen zur Angriffserkennung in deutschen Krankenhäusern auch nach dieser Aufarbeitung erst am Anfang. Die erarbeiteten Vorschläge sollten im Rahmen der Fortschreibung des B3S intensiv unter den Betreibern diskutiert werden. Auch das BSI sollte eingebunden werden. Dabei muss das letztendliche Zielniveau für SzA in der medizinischen Versorgung

festgelegt werden. Der vorliegende Bericht stellt eine Momentaufnahme dar und es ist zu erwarten, dass sich das Umfeld für IT-Sicherheit in Kritischen Infrastrukturen wie Krankenhäusern sowohl regulatorisch (durch die Umsetzung der NIS2-Richtlinie) als auch technisch weiterentwickeln wird. Darüber hinaus gewinnt das Thema international an Bedeutung. Es sollten daher internationale Entwicklungen beobachtet und, soweit passend, im weiteren Verlauf berücksichtigt werden. Das Projektteam regt an, die Ergebnisse der B3S-Aktivitäten „Medizinische Versorgung" in Deutschland durch englische Übersetzungen in die internationale Debatte einzubringen. Auch die Ergebnisse dieses Projekts sollten sinnvoll in die englischsprachige Community integriert werden, um den internationalen Austausch zu fördern.

Anlage 1: Literaturverzeichnis

[1] *Zweites Gesetz zur Erhöhung der Sicherheit informationstechnischer Systeme*. 2021, S. 1122. Zugegriffen: 20. Dezember 2021. [Online]. Verfügbar unter: http://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl121s1122.pdf

[2] *Gesetz zur Erhöhung der Sicherheit informationstechnischer Systeme*. 2015, S. 1324–1331. [Online]. Verfügbar unter: http://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl115s1324.pdf

[3] „Orientierungshilfe zum Einsatz von Systemen zur Angriffserkennung (Version 1.0)". 26. September 2022. [Online]. Verfügbar unter: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/KRITIS/oh-sza.pdf

[4] Deutsche Krankenhausgesellschaft, „Branchenspezifischer Sicherheitsstandard ‚Medizinische Versorgung'". 8. Dezember 2022. [Online]. Verfügbar unter: https://www.dkgev.de/fileadmin/default/Mediapool/2_Themen/2.1_Digitalisierung_Daten/2.1.4._IT-Sicherheit_und_technischer_Datenschutz/2.1.4.1._IT-Sicherheit_im_Krankenhaus/Branchenspezifischer_Sicherheitsstandard_Medizinische_Versorgung_v1.2_Stand_2022-12-08.pdf

[5] B. Jungclaus, „Detektion von Ransomware in der IT-Systemlandschaft bei Betreibern Kritischer Infrastruktur", 2023, doi: 10.25933/OPUS4-2903.

[6] Volker Amelung, Mike Angelkorte, Boris Augurzky, Robert Brauer, Felix Freigang, Frank Fritzsche, Alexander Geissler, Aydan Göller, Alexander Haering, Malte Haring, Johannes Hollenbach, Manuel Luckmann, Kerstin Materne, Ronan O'Connor, Jens Peukert, Franziska Püschner, Lorenz von Roehl Armin Scheuer, Anne Snowdon, Christoph Steuber, Sylvia Thun, Isa-bel Vollrath, und Anne Wiesmann, „Zwischenbericht DigitalRadar - Ergebnisse der ersten nationalen Reifegradmessung deutscher Krankenhäuser", 2022. [Online]. Verfügbar unter: https://www.digitalradar-krankenhaus.de/download/220914_Zwischenbericht_DigitalRadar_Krankenhaus.pdf

[7] Statistisches Bundesamt, „Grunddaten der Krankenhäuser". 7. April 2022. [Online]. Verfügbar unter: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Krankenhaeuser/Publikationen/Downloads-Krankenhaeuser/grunddaten-krankenhaeuser-2120611207004.html

[8] Bundesministerium des Innern, *Erste Verordnung zur Änderung der BSI-Kritisverordnung*. 2017. [Online]. Verfügbar unter: https://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl117s1903.pdf

[9] Bundesamt für Sicherheit in der Informationstechnik, „Die Lage der IT-Sicherheit in Deutschland 2020". 20. Oktober 2020. [Online]. Verfügbar unter: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Lageberichte/Lagebericht2020.pdf

[10] INFO GmbH Markt- und Meinungsforschung, „Untersuchung zur Wirksamkeit der IT-Sicherheitsgesetze unter Betreibern Kritischer Infrastrukturen - Ergebnisbericht", Apr. 2023. [Online]. Verfügbar unter: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/KRITIS/evaluierung-itsig2-ergebnisbericht.pdf

[11] *Gesetz zum Schutz elektronischer Patientendaten in der Telematikinfrastruktur*. 2020. [Online]. Verfügbar unter: https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/Gesetze_und_Verordnungen/GuV/P/PDSG_bgbl.pdf

[12] Bundesregierung, *Gesetz zur Beschleunigung der Digitalisierung des Gesundheitswesens*. 2024. [Online]. Verfügbar unter: https://www.recht.bund.de/bgbl/1/2024/101/VO.html

[13] Bundesministerium des Innern, *Entwurf eines Gesetzes zur Umsetzung der CER-Richtlinie und zur Stärkung der Resilienz kritischer Anlagen*. 2023. [Online]. Verfügbar unter: https://www.bmi.bund.de/SharedDocs/gesetzgebungsverfahren/DE/Downloads/referentenentwuerfe/KM4/KRITIS-DachG-1.pdf

[14] Bundesministerium des Innern, *Entwurf eines Gesetzes zur Umsetzung der CER-Richtlinie und zur Stärkung der Resilienz kritischer Anlagen*. 2023. [Online]. Verfügbar unter:

https://www.bmi.bund.de/SharedDocs/gesetzgebungsverfahren/DE/Downloads/referentenentwuerfe/KM4/KRITIS-DachG-2.pdf
[15] Bundesministeriums des Innern und für Heimat, *Entwurf eines Gesetzes zur Umsetzung der NIS-2-Richtlinie und zur Regelung wesentlicher Grundzüge des Informationssicherheitsmanagements in der Bundesverwaltung*. 2024. Zugegriffen: 8. Mai 2024. [Online]. Verfügbar unter: https://www.bmi.bund.de/SharedDocs/gesetzgebungsverfahren/DE/Downloads/referentenentwuerfe/CI1/NIS-2-RefE.pdf
[16] Bundesamt für Sicherheit in der Informationstechnik, „Baustein OPS.1.1.5 Protokollierung (IT-Grundschutz-Kompendium)". [Online]. Verfügbar unter: https://www.bsi.bund.de/DE/Themen/ITGrundschutz/ITGrundschutzKompendium/bausteine/OPS/OPS_1_1_5_Protokollierung.html
[17] Bundesamt für Sicherheit in der Informationstechnik, „DER.1 Detektion von sicherheitsrelevanten Ereignissen (IT-Grundschutz-Baustein)". Edition 2023. [Online]. Verfügbar unter: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Grundschutz/IT-GS-Kompendium_Einzel_PDFs_2023/05_DER_Detektion_und_Reaktion/DER_1_Detektion_von_sicherheitsrelevanten_Ereignissen_Edition_2023.pdf
[18] Bundesamt für Sicherheit in der Informationstechnik, „DER.2.1 Behandlung von Sicherheitsvorfällen (IT-Grundschutz-Baustein)". Edition 2023. [Online]. Verfügbar unter: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Grundschutz/IT-GS-Kompendium_Einzel_PDFs_2023/05_DER_Detektion_und_Reaktion/DER_2_1_Behandlung_von_Sicherheitsvorfaellen_Edition_2023.pdf
[19] Bundesregierung, *Entwurf eines Gesetzes zur Erhöhung der Sicherheit informationstechnischer Systeme*. 2015. [Online]. Verfügbar unter: https://dserver.bundestag.de/btd/18/040/1804096.pdf
[20] „Bekanntmachung des Handbuchs der Rechtsförmlichkeit". Bundesanzeiger, 22. September 2008. [Online]. Verfügbar unter: https://www.bmj.de/SharedDocs/Publikationen/DE/Fachpublikationen/Handbuch_der_Rechtsfoermlichkeit.pdf
[21] Bundesverband IT-Sicherheit e.V. - TeleTrust-Arbeitskreis „Stand der Technik", „IT-Sicherheitsgesetz und Datenschuzt-Grundversorgung: Handreichnung zum ‚Stand der Technik' Technische und Organisatorische Maßnahmen". Mai 2023. [Online]. Verfügbar unter: https://www.teletrust.de/publikationen/broschueren/stand-der-technik/?tx_reintdownloadmanager_reintdlm%5Bdownloaduid%5D=11375&cHash=911131cf4407f73e8649e9ed5f512c6e
[22] Bundesamt für Sicherheit in der Informationstechnik, „KRITIS-Sektor Gesundheit: Informationssicherheit in der stationären medizinischen Versorgung Rahmenbedingungen, Status Quo, Handlungsfelder - Ergebnisse einer qualitativen Studie". 30. Juni 2020. [Online]. Verfügbar unter: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Studien/KRITIS/Studie_Informationssicherheit_stationaere_med_Versorgung.pdf
[23] ISO/IEC, „ISO/IEC 27033-1: Information technology - Security techniques - Network security - Part 1: Overview and concepts". 15. August 2015.
[24] ISO/IEC, „ISO/IEC 27033-2: Information technology - Security techniques - Network security - Part 2: Guidelines for the design and implementation of network security". 15. August 2012.
[25] ISO/IEC, „ISO/IEC 27035-1: Information technology - Information security incident management - Part 1: Principles and process". Februar 2023.
[26] ISO/IEC, „ISO/IEC 27035-2: Information technology - Information security incident management - Part 2: Guidelines to plan and prepare for incident response". Februar 2023.
[27] ISO/IEC, „ISO/IEC 27035-3: Information technology - Information security incident management - Part 3: Guidelines for ICT incident response operations". September 2020.
[28] *IEC 62443-3-3: Industrial communication networks – Network and system security – Part 3-3: System security requirements and security levels*, Edition 1.0, 2013-08. in International standard / IEC, no. 62443-3–3. Geneva: IEC Central Office, 2013.
[29] *IEC 62443-4-2: Security for industrial automation and control systems. Technical security requirements for IACS components*, Edition 1.0, 2019-02. in International standard / IEC, no. 62443-4–2. Geneva: IEC Central Office, 2019.

[30] ISO/IEC, „ISO/IEC 27039: Information technology — Security techniques — Selection, deployment and operations of intrusion detection and prevention systems (IDPS)". 1. Mai 2016. [Online]. Verfügbar unter: https://www.beuth.de/de/norm/iso-iec-27039/231353426
[31] K. A. Scarfone und P. M. Mell, „Guide to Intrusion Detection and Prevention Systems (IDPS)", National Institute of Standards and Technology, Gaithersburg, MD, NIST SP 800-94, 2007. doi: 10.6028/NIST.SP.800-94.
[32] National Institute of Standards and Technology (NIST), „Computer Security Incident Handling Guide, Recommendations of the National Institute of Standards and Technology, Special Publication 800-61, Revision 2". August 2012. [Online]. Verfügbar unter: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-61r2.pdf
[33] National Institute of Standards and Technology, „Framework for Improving Critical Infrastructure Cybersecurity". 16. April 2018. doi: https://doi.org/10.6028/NIST.CSWP.04162018.
[34] Bundesamt für Sicherheit in der Informationstechnik, „Mindeststandard des BSI zur Protokollierung und Detektion von Cyber-Angriffen Version 2.0". 29. Juni 2023. [Online]. Verfügbar unter: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Mindeststandards/Mindeststandard_BSI_Protokollierung_und_Detektion_Version_2_0.pdf?__blob=publicationFile&v=3
[35] „Informationssicherheit, Cybersicherheit und Schutz der Privatsphäre – Informationssicherheitsmaßnahmen (ISO/IEC 27002:2022); Deutsche Fassung EN ISO/IEC 27002:2022". Januar 2024.
[36] National Institute of Standards and Technology, „The NIST Cybersecurity Framework (CSF) 2.0", National Institute of Standards and Technology, Gaithersburg, MD, NIST CSWP 29, Feb. 2024. doi: 10.6028/NIST.CSWP.29.
[37] Deutsches Institut für Normung, „Anwendung des Risikomanagements für IT-Netzwerke, die Medizinprodukte beinhalten - Teil 1: Aufgaben, Verantwortlichkeiten und Aktivitäten". 2011. [Online]. Verfügbar unter: https://www.beuth.de/de/norm/din-en-80001-1/145057440
[38] IEC, „IEC TR 80001-2-2:2012 Application of risk management for IT-networks incorporating medical devices - Part 2-2: Guidance for the disclosure and communication of medical device security needs, risks and controls". Juli 2012. [Online]. Verfügbar unter: https://www.vde-verlag.de/iec-normen/218983/iec-tr-80001-2-2-2012.html
[39] IHE International, „Audit Trail and Node Authentication (ATNA) Profile in IHE IT Infrastructure (ITI) Technical Framework, Volume 1". 2023. Zugegriffen: 30. Mai 2024. [Online]. Verfügbar unter: https://profiles.ihe.net/ITI/TF/Volume1/ch-9.html
[40] IHE International, „Record Audit Event [ITI-20] in IHE IT Infrastructure (ITI) Technical Framework, Volume 2". 2023. Zugegriffen: 30. Mai 2024. [Online]. Verfügbar unter: https://profiles.ihe.net/ITI/TF/Volume2/ITI-20.html#3.20
[41] „DIN EN ISO 27799:2016-12, Medizinische Informatik_- Informationssicherheitsmanagement im Gesundheitswesen bei Verwendung der ISO/IEC_27002 (ISO_27799:2016); Englische Fassung EN_ISO_27799:2016", Beuth Verlag GmbH. doi: 10.31030/2561354.
[42] J. Marron, „Implementing the Health Insurance Portability and Accountability Act (HIPAA) Security Rule: A Cybersecurity Resource Guide", National Institute of Standards and Technology, Gaithersburg, MD, NIST SP 800-66r2, 2024. doi: 10.6028/NIST.SP.800-66r2.
[43] U.S. Department of Health & Human Services, „Health Care and Public Health Sector Cybersecurity Framework Implementation Guide". März 2023. [Online]. Verfügbar unter: https://aspr.hhs.gov/cip/hph-cybersecurity-framework-implementation-guide/Pages/Foreword.aspx
[44] Healthcare & Public Health Sector Coordinating Council, „Health Industry Cybersecurity Practices: Managing Threats and Protecting Patients". 2023. [Online]. Verfügbar unter: https://405d.hhs.gov/Documents/HICP-Main-508.pdf
[45] Healthcare & Public Health Sector Coordinating Council, „Technical Volume 2: Cybersecurity Practices for Medium and Large Healthcare Organizations". 2023. [Online]. Verfügbar unter: https://405d.hhs.gov/Documents/tech-vol2-508.pdf
[46] Healthcare & Public Health Sector Coordinating Council, „Technical Volume 1: Cybersecurity Practices for Small Healthcare Organizations". 2023. [Online]. Verfügbar unter: https://405d.hhs.gov/Documents/tech-vol1-508.pdf
[47] „Branchenspezifischer Sicherheitsstandard für die Verteilung von Fernwärme". 15. Februar 2021. Zugegriffen: 15. Dezember 2023. [Online]. Verfügbar unter: https://www.bdew.de/energie/b3s-fernwaermenetze/

[48] DVGW Technischen Komitee „IT-Sicherheit" sowie DWA-Arbeitsgruppe „Cyber-Sicherheit", „Merkblatt W 1060 2022-04 Wasser - IT-Sicherheit – Branchenstandard Wasser/Abwasser". [Online]. Verfügbar unter: https://www.dvgw.de/leistungen/publikationen/publikationsliste/it-sicherheitsstandard-wasser-abwasser-b3s

[49] „DIN 13080 Gliederung des Krankenhauses in Funktionsbereiche und Funktionsstellen". 2016. Zugegriffen: 28. Februar 2024. [Online]. Verfügbar unter: https://www.din.de/de/mitwirken/normenausschuesse/nabau/veroeffentlichungen/wdc-beuth:din21:252635669

[50] Kathryn Knerler, Ingrid Parker, und Carson Zimmerman, *11 Strategies of a World-Class Cybersecurity Operations Center*. MITRE. [Online]. Verfügbar unter: https://www.mitre.org/sites/default/files/2022-04/11-strategies-of-a-world-class-cybersecurity-operations-center.pdf

[51] European Union Agency for Cybersecurity, „ENISA threat landscape 2023", Okt. 2023. [Online]. Verfügbar unter: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2023

[52] Die Bundesregierung, „Antwort der Bundesregierung auf die Kleine Anfrage der Fraktion der CDU/CSU – Drucksache 20/10657 – Cyberattacken auf Krankenhäuser – Sachstand und Handlungsbedarf". 3. April 2024. Zugegriffen: 11. April 2024. [Online]. Verfügbar unter: https://dserver.bundestag.de/btd/20/109/2010907.pdf

[53] Bundesamt für Sicherheit in der Informationstechnik, „Mindeststandard des BSI zur Protokollierung und Detektion von Cyber-Angriffen nach § 8 Absatz 1 Satz 1 BSIG – Version 1.0a vom 25.02.2021". 25. Februar 2021. Zugegriffen: 23. Mai 2024. [Online]. Verfügbar unter: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Mindeststandards/Mindeststandard_BSI_Protokollierung_und_Detektion_Version_1_0a.pdf?__blob=publicationFile&v=5

[54] Der Bundesbeauftragte für den Datenschutz und die Informationsfreiheit, „Stellungnahme des Bundesbeauftragten für den Datenschutz und die Informationsfreiheit zum Entwurf eines Zweiten Gesetzes zur Erhöhung der Sicherheit informationstechnischer Systeme (IT-Sicherheitsgesetz 2.0)". 18. Dezember 2020. Zugegriffen: 23. Mai 2024. [Online]. Verfügbar unter: https://www.bfdi.bund.de/SharedDocs/Downloads/DE/DokumenteBfDI/Stellungnahmen/2020/StgN_IT-Sicherheitsgesetz-2.0.pdf?__blob=publicationFile&v=4

[55] CROWDSTRIKE, „CROWDSTRIKE SERVICES CYBER FRONT LINES REPORT - INCIDENT RESPONSE AND PROACTIVE SERVICES FROM 2020 AND INSIGHTS THAT MATTER FOR 2021", 2020. Zugegriffen: 23. Mai 2024. [Online]. Verfügbar unter: https://www.crowdstrike.com/resources/reports/cyber-front-lines/

[56] Mandiant, „M-Trends2024 Special Report", 2024. Zugegriffen: 23. Mai 2024. [Online]. Verfügbar unter: https://www.mandiant.com/m-trends

Anlage 2: Fragebogen Bestandsaufnahme Umsetzung SzA

1. Einleitung

Die Erhebung mit diesem Fragebogen richtet sich an Betreiber von Krankenhäusern in Deutschland. Sie wird durchgeführt, um einen Überblick über den aktuellen Umsetzungsstand bzgl. Systeme zur Angriffserkennung in deutschen Krankenhäusern zu erhalten. Die Durchführung erfolgt von der Technischen Hochschule Brandenburg (THB) im Rahmen einer Unterstützungsleistung, die von der Deutschen Krankenhausgesellschaft im Zusammenhang mit der Fortschreibung des Branchenspezifischen Sicherheitsstandards (B3S) für die Medizinische Versorgung beauftragt wurde. Die Inhalte des Fragebogens speisen sich aus der Orientierungshilfe SzA (OH SzA) des BSI, welche im Rahmen der THB-Unterstützung für den Betrieb von Krankenhäusern konkretisiert werden soll. Ergebnisse aus der Umfrage werden verwendet, um den B3S hinsichtlich der Anteile SzA fortzuschreiben. Eine fundierte Basis zum aktuellen Umsetzungsstatus ist essenziell für die Realitätsnähe des zukünftigen B3S - nehmen Sie sich daher bitte bewusst die Zeit für die Beantwortung der Fragen und insb. auch der Freittextfelder für weitere Informationen.

Die Ergebnisse der Erhebung werden Ende August in einem ersten Workshop bei der DKG in Berlin mit Vertretern aus der Branche diskutiert und ausgewertet werden. Sofern Sie Interesse an der Mitwirkung haben, können Sie sich gern bei uns melden (medsec@th-brandenburg.de). Die Einladungen erfolgen sodann noch mit eigener E-Mail. Im weiteren Verlauf der Unterstützung erfolgt die Information primär über den UP KRITIS mit seinem Branchenarbeitskreis Medizinische Versorgung und den entsprechenden Unterarbeitskreisen zum Thema.

Die Erhebung erfolgt anonymisiert und die Ergebnisse der Erhebung dienen als Grundlage für die weitere Analyse; Details mit Rückschlussmöglichkeit auf einzelne Betreiber werden nicht veröffentlicht (Weitergabe nur an Auftraggeber). Abgeleitete Ergebnisse ohne Rückschlussmöglichkeit auf einzelne Antworten werden genutzt und ggf. auch veröffentlicht werden.

Wichtig: Leider bietet das Umfragetool evasys NICHT die Möglichkeit, den Fragebogen mit Ihren Antworten am Ende für Ihre Akten zu exportieren. Bitte notieren Sie sich deshalb Ihre Antworten gleich beim Ausfüllen mit, wenn Sie diese benötigen.

Nach vorläufiger Einschätzung der THB unterscheidet sich der Umsetzungsstand bzgl. SzA innerhalb von Organisationen in den unterschiedlichen Bereichen. Daher wird in einzelnen Fragen eine Differenzierung vorgenommen. Hierfür werden folgende Netzbereiche herangezogen:

IT – Informationstechnik: inkl. Kommunikationstechnik (Abschnitte 4.5.1 sowie 4.5.2 des B3S MV)

MT – Medizintechnik: bezieht sich analog zu DIN 80001 auf "IT-Netzwerke, die Medizinprodukte enthalten"

VT – Versorgungstechnik: inkl. Gebäude-Leit-Technik (GLT) (Abschnitt 4.5.3 des B3S MV).

2. Informationen zur antwortenden Organisation

Die Antwortmöglichkeiten orientieren sich weitgehend am ersten Zwischenbericht zum Digitalradar.

2.1 Ihre Organisation betreibt mindestens eine Anlage in der stationären medizinischen Versorgung, welches die Kriterien lt. BSI- Gesetz / BSI-KritisV erfüllt und somit unter die Definition einer Kritischen Infrastruktur fällt.

- Ja,
- Nein

2.2 Welche Trägerschaft trifft auf Ihr Krankenhaus/Ihre Krankenhäuser zu?

- Privat,
- Öffentlich,
- Freigemeinnützig,
- Sonstiges (Bitte spezifizieren)

2.3 Welche?
(Freitext)

2.4 Ihre Organisation betreibt eine Krankenhauskapazität an Betten:

- 800 und mehr,
- 600 – 799,
- 500 – 599,
- 400 – 499,
- 300 – 399,
- 200 – 299,
- 100 – 199,
- 0 – 99

2.5 Wie viele Mitarbeitende arbeiten insgesamt in Ihrer Organisation (Vollzeit-Äquivalente)?
(Freitext)

2.6 Auf wie viele Krankenhausstandorte gemäß §293 Abs. 6 SGB V bezieht sich die Beantwortung dieses Fragebogens?
(Freitext)

3. Eigene Einschätzung zum Reifegrad SzA

Ordnen Sie bitte den aktuellen Stand der SzA-Einführung in Ihrer Organisation entlang der Umsetzungsgrade aus der Orientierungshilfe SzA des BSI (Abschnitt 4 Nachweis von Systemen zur

Angriffserkennung) zu; der Begriff "Bereich" bezieht sich gemäß OH SzA auf a) Protokollierung, b) Detektion und c) Reaktion mit den jeweils in Kapitel 3 formulierten Anforderungen.

3.1 Angriffserkennung in der Organisation insgesamt

- 0 - Keine Maßnahmen / keine Planungen
- 1 - Planungen vorhanden / mind. 1 Bereich ohne Umsetzungen
- 2 - alle Bereiche begonnen / offene MUSS- Anforderungen
- 3 - alle MUSS- Anforderungen für alle Bereiche / KVP mindestens in Planung
- 4 - alle MUSS- Anforderungen erfüllt / alle SOLL- Anforderungen erfüllt oder stichhaltig begründet ausgeschlossen / KVP etabliert
- 5 - wie 4, jedoch Umsetzung zusätzlicher Maßnahmen aus der Risikoanalyse

3.2 Angriffserkennung in der IT

- 0 - Keine Maßnahmen / keine Planungen
- 1 - Planungen vorhanden / mind. 1 Bereich ohne Umsetzungen
- 2 - alle Bereiche begonnen / offene MUSS- Anforderungen
- 3 - alle MUSS- Anforderungen für alle Bereiche / KVP mindestens in Planung
- 4 - alle MUSS- Anforderungen erfüllt / alle SOLL- Anforderungen erfüllt oder stichhaltig begründet ausgeschlossen / KVP etabliert
- 5 - wie 4, jedoch Umsetzung zusätzlicher Maßnahmen aus der Risikoanalyse

3.3 Angriffserkennung in der MT

- 0 - Keine Maßnahmen / keine Planungen
- 1 - Planungen vorhanden / mind. 1 Bereich ohne Umsetzungen
- 2 - alle Bereiche begonnen / offene MUSS- Anforderungen
- 3 - alle MUSS- Anforderungen für alle Bereiche / KVP mindestens in Planung
- 4 - alle MUSS- Anforderungen erfüllt / alle SOLL- Anforderungen erfüllt oder stichhaltig begründet ausgeschlossen / KVP etabliert
- 5 - wie 4, jedoch Umsetzung zusätzlicher Maßnahmen aus der Risikoanalyse

3.4 Angriffserkennung in der VT

- 0 - Keine Maßnahmen / keine Planungen
- 1 - Planungen vorhanden / mind. 1 Bereich ohne Umsetzungen
- 2 - alle Bereiche begonnen / offene MUSS- Anforderungen
- 3 - alle MUSS- Anforderungen für alle Bereiche / KVP mindestens in Planung
- 4 - alle MUSS- Anforderungen erfüllt / alle SOLL- Anforderungen erfüllt oder stichhaltig begründet ausgeschlossen / KVP etabliert
- 5 - wie 4, jedoch Umsetzung zusätzlicher Maßnahmen aus der Risikoanalyse

4. Branchenspezifische Risikolage inkl. branchenspezifischer Bedrohungen

Sofern Ihr Risiko-Management keine entsprechenden Auswertungen ermöglicht, beantworten Sie die Fragen bitte nach bestem Wissen und Gewissen (Experteneinschätzung).

4.1 Welche 5 Gefährdungen der Informationssicherheit sind aktuell die Top-Risiken für Ihre Organisation?

- Nichtverfügbarkeit wichtiger, medizinisch relevanter Daten im Diagnose-Prozess
- Nichtverfügbarkeit wichtiger medizinisch relevanter Daten im Entlassungs-Prozess
- Nichtverfügbarkeit von behandlungsrelevanten Logistikketten
- Manipulation von medizinisch relevanten Daten im Diagnose- Prozess
- Manipulation von medizinisch relevanten Daten im Entlassungs-Prozess
- Verlust der Datenauthentizität
- Fremdsteuerung/Manipulation von relevanten Infrastrukturkomponenten
- Nichtverfügbarkeit wichtiger medizinisch relevanter Daten im Therapie-Prozess
- Nichtverfügbarkeit von für den Behandlungsprozess wichtiger Prozess- und Freigabeinformationen
- Inkonsistenzen in für den Behandlungsprozess relevanten Datenbeständen
- Manipulation von medizinisch relevanten Daten im Therapie- Prozess
- Unterbrechung von behandlungsrelevanten Kommunikationsabläufen
- Fremdsteuerung/Manipulation von medizinische relevanten IT-Systemen
- Nichtverfügbarkeit wichtiger medizinisch relevanter Daten im Pflege-Prozess
- Nichtverfügbarkeit von behandlungsprozessrelevanten IT-Systemen
- Inkonsistenzen bei der Übertragung von für den Behandlungsprozess relevanten Datenbeständen
- Manipulation von medizinisch relevanten Daten im Pflege-Prozess
- Vertraulichkeitsverlust bei besonders sensiblen Patienten- und Behandlungsinformationen
- Fremdsteuerung/Manipulation von Medizingeräten
- Sonstige (bitte spezifizieren)

4.2 Welche?

(Freitext)

4.3 Welche 5 Gefährdungen der Informationssicherheit sind im Risikopotential für Ihre Organisation besonders "aufgestiegen" im Jahr 2022 / 2023?

- Nichtverfügbarkeit wichtiger, medizinisch relevanter Daten im Diagnose-Prozess
- Nichtverfügbarkeit wichtiger medizinisch relevanter Daten im Entlassungs-Prozess
- Nichtverfügbarkeit von behandlungsrelevanten Logistikketten

- Manipulation von medizinisch relevanten Daten im Diagnose- Prozess
- Manipulation von medizinisch relevanten Daten im Entlassungs-Prozess
- Verlust der Datenauthentizität
- Fremdsteuerung/Manipulation von relevanten Infrastrukturkomponenten
- Nichtverfügbarkeit wichtiger medizinisch relevanter Daten im Therapie-Prozess
- Nichtverfügbarkeit von für den Behandlungsprozess wichtiger Prozess- und Freigabeinformationen
- Inkonsistenzen in für den Behandlungsprozess relevanten Datenbeständen
- Manipulation von medizinisch relevanten Daten im Therapie- Prozess
- Unterbrechung von behandlungsrelevanten Kommunikationsabläufen
- Fremdsteuerung/Manipulation von medizinische relevanten IT-Systemen
- Nichtverfügbarkeit wichtiger medizinisch relevanter Daten im Pflege-Prozess
- Nichtverfügbarkeit von behandlungsprozessrelevanten IT-Systemen
- Inkonsistenzen bei der Übertragung von für den Behandlungsprozess relevanten Datenbeständen
- Manipulation von medizinisch relevanten Daten im Pflege-Prozess
- Vertraulichkeitsverlust bei besonders sensiblen Patienten- und Behandlungsinformationen
- Fremdsteuerung/Manipulation von Medizingeräten
- Sonstige (bitte spezifizieren)

4.4 Welche?
(Freitext)

4.5 Bitte kennzeichnen Sie diejenigen Bedrohungen aus dem B3S (Abschnitt 4.1), welche Sie für Ihre Organisation insgesamt für relevant halten.

- Höhere Gewalt und Elementarschadensereignisse
- Manipulation, Diebstahl, Verlust, Zerstörung von IT oder IT-relevanten Anlagen und Anlagenteilen
- Hacking und Manipulation
- Gezielte Störung / Verhinderung von Diensten, z. B. distributed denial of service (DDoS), gezielte Systemabstürze, u. ä.
- Advanced Persistent Threat (APT)
- Abhängigkeiten von Dienstleistern und Herstellern (Ausfall externer Dienstleister, unberechtigter Zugriff, versteckte Funktionen in Hard- und Software)
- Beschädigung oder Zerstörung verfahrenstechnischer Komponenten, Ausrüstungen und Systeme
- Schadprogramme / Ransomware Social Engineering
- E-Mail-Account-Übernahme / Spamming
- Ausfall von Basisinfrastrukturen mit direktem Bezug zur IT (Sekundäreffekte, z. B. Strom und TK)
- Terroristische Akte (physisch mit Wirkung auf die IT oder direkt IT- bezogen)

- Systemmissbrauch (Innentäter) und unbefugter Zugriff
- Identitätsmissbrauch (Phishing, Skimming, Zertifikatsfälschung)
- Sonstige (bitte spezifizieren)

4.6 Welche?
(Freitext)

4.7 Bitte kennzeichnen Sie diejenigen Bedrohungen aus dem B3S (Abschnitt 4.1), welche Sie für die Angriffserkennung (SzA) in Ihrer Organisation heranziehen.

- Höhere Gewalt und Elementarschadensereignisse
- Manipulation, Diebstahl, Verlust, Zerstörung von IT oder IT-relevanten Anlagen und Anlagenteilen
- Hacking und Manipulation
- Gezielte Störung / Verhinderung von Diensten, z. B. distributed denial of service (DDoS), gezielte Systemabstürze, u. ä.
- Advanced Persistent Threat (APT)
- Abhängigkeiten von Dienstleistern und Herstellern (Ausfall externer Dienstleister, unberechtigter Zugriff, versteckte Funktionen in Hard- und Software)
- Beschädigung oder Zerstörung verfahrenstechnischer Komponenten, Ausrüstungen und Systeme
- Schadprogramme / Ransomware Social Engineering
- E-Mail-Account-Übernahme / Spamming
- Ausfall von Basisinfrastrukturen mit direktem Bezug zur IT (Sekundäreffekte, z. B. Strom und TK)
- Terroristische Akte (physisch mit Wirkung auf die IT oder direkt IT- bezogen)
- Systemmissbrauch (Innentäter) und unbefugter Zugriff
- Identitätsmissbrauch (Phishing, Skimming, Zertifikatsfälschung)
- Sonstige (bitte spezifizieren)

4.8 Welche?
(Freitext)

4.9 Bitte spezifizieren Sie, aus welchen Quellen Sie Bedrohungsinformationen zur Informationssicherheit (Cyber Threat Intelligence) ziehen. Bitte nutzen Sie dringend die Freitext-Option und hinterlegen Sie so konkret wie möglich (Anbieter, Produkt, URL) Ihre Quellen.

- BSI Website CIRCL
- Collaborative Research Into Threats (CRITs)
- ElecticIQ
- Malware Information Sharing Platform (MISP); allgemein - bitte spezifizieren
- Threat Connect Vipre ThreatIQ

- BSI MISP abuse.ch
- Collective Intelligence Framework (CIF)
- Facebook Threat Exchange (TX) Open Threat Exchange (OTX)
- ThreatQ
- Sonstige (bitte spezifizieren)

4.10 Welche?
(Freitext)

4.11 Bitte spezifizieren Sie, aus welchen Quellen Sie Informationen zu aktuellen Angriffsmustern für technische Vulnerabilitäten beziehen (Auswahloptionen gemäß OH SzA)

- Hersteller (Hard- und Software)
- Behörden
- Medien
- weitere relevante Stellen (bitte spezifizieren)

4.12 Welche?
(Freitext)

4.13 Zum Aufbau und zum Betrieb von SzA wird regelmäßig die Verwendung von Frameworks (z. B. MITRE ATT&CK) empfohlen. Falls Sie ebenfalls ein solches Framework herangezogen haben, wofür?

- Zum Aufbau der Detektionsarchitektur
- Zur Visualisierung und Einordnung erkannter Angriffe
- Zur Validierung der Abdeckung der SzA
- Zur Überprüfung der eingesetzten Regelsätze
- Bisher nicht verwendet, aber geplant
- Nicht verwendet und auch nicht geplant
- Andere (bitte spezifizieren)

4.14 Welche?
(Freitext)

5. Branchenspezifische Prozesseigenheiten, Abhängigkeiten

5.1 Bitte kennzeichnen Sie diejenigen Geschäftsprozesse aus dem B3S, welche Sie für Ihre Organisation hinsichtlich SzA für relevant halten.

- Vorbereitung/Aufnahme
- Diagnostik
- Unterbringung und Pflege

- Therapie
- Entlassung

5.2 Welche weiteren Prozesse (außerhalb der im B3S benannten) sind in Ihrer Organisation relevant?
(Freitext)

5.3 Welchen Abdeckungsgrad bzgl. Ihrer relevanten Geschäftsprozesse erreicht Ihre Business Impact Analyse ungefähr?

- <10%
- 10% - 25%
- 25% - 50%
- 50% - 75%
- 75% - 90%
- >90%

5.4 Bezieht Ihre Organisation für die kDL relevante IT-Dienste aus der Cloud? Bitte beschreiben Sie diese.
(Freitext)

6. IT-Systemarchitektur: Branchenspezifische Strukturanalyse inkl. Netzplan

6.1 Wird in Ihrem Haus eine Inventarisierungssoftware für die betriebene IT/ Versorgungstechnik/MT eingesetzt? Wie gut bildet diese Inventarisierungssoftware aus Ihrer Sicht die wirklichen Verhältnisse ab?

- Ja, alle betriebenen Geräte sind abgedeckt.
- Ja, die meisten Geräte werden erfasst.
- Ja, aber es gibt viele Geräte, die nicht erfasst werden können.
- Nein, aber der Einsatz einer solchen Software ist geplant.
- Nein, und der Einsatz ist auch NICHT geplant.

6.2 Wie viele (gerundet) Medizingeräte befinden sich in Ihrer IT-Infrastruktur?
(Freitext)

6.3 Wie viele (gerundet) netzwerkfähige Medizingeräte befinden sich in Ihrer IT-Infrastruktur?
(Freitext)

6.4 Wie viele (gerundet) Zielobjekte (Assets) befinden sich in Ihrer IT-Infrastruktur inkl. Netzwerktechnik und Kommunikationstechnik?
(Freitext)

6.5 Wie viele (gerundet) Zielobjekte (Assets) befinden sich in der Versorgungstechnik (VT) inkl. GLT?
(Freitext)

6.6 Wie viele (gerundet) Server (inkl. virtuelle Server) befinden sich in Ihrer IT-Infrastruktur?
(Freitext)

6.7 Wie viele Netzsegmente (VLANs) weist Ihre IT-Infrastruktur auf?
(Freitext)

6.8 Nach welcher Gruppierung erfolgt die Netzsegmentierung?

- Organisationseinheiten
- MT vs. IT vs. VT
- andere (bitte spezifizieren)

6.9 Welche?
(Freitext)

6.10 Wie hart sind die einzelnen Netzsegmente voneinander getrennt?

- Geräte dürfen nur im eigenen Netzsegment kommunizieren
- Einzelne, festgelegte Geräte dürfen über Segmentgrenzen hinweg kommunizieren
- Nur festgelegte Geräte dürfen NICHT über Segmentgrenzen hinweg kommunizieren.
- Beschränkungen zwischen Segmenten gibt es bisher nicht, sie sind aber geplant
- Beschränkungen zwischen Segmenten gibt es bisher nicht, sie sind auch NICHT geplant

6.11 Inwieweit werden Protokollierungsdaten zentral eingesammelt?
lokale Vorhaltung

- Zentral für einzelne Organisationseinheiten
- Zentral für ein ganzes Haus
- Zentral für die Betreiberorganisation
- Zentral und organisationsübergreifend

6.12 Welche Anzahl an Systemen ist in die zentrale Protokollierung einbezogen?
(Freitext)

6.13 Auf Basis welcher Merkmale wird die Reihenfolge zur Einbeziehung von Systemen festgelegt?

- Zuordnung IT, MT, VT
- Relevanz des Systems für kDL
- Kritikalität der Systeme

- Betriebssystem
- Netzgrenzen zuerst
- andere (bitte spezifizieren)

6.14 Welche?
(Freitext)

6.15 Wie viele (gerundet) Datensätze werden durch die zentrale Protokollierung täglich eingesammelt?
(Freitext)

6.16 Wie viele (gerundet) qualifizierte Events haben Sie täglich?
(Freitext)

7. Übersicht über eingesetzte Detektions-Technologien und -services

7.1 Welche Technologien kommen zur Angriffserkennung zum Einsatz?

- HIDS / HIPS
- NIDS / NIPS
- EDR
- XDR
- SIEM
- Log-Aggregator
- Mail-Filtering
- Honeypots/Deception Systeme
- (NG) Firewalls
- Antivirus mit zentralem Management
- Web-Reputationfilter (Web-Washer)
- Andere (bitte spezifizieren)

7.2 Welche?
(Freitext)

7.3 Haben Sie NIDS an den Netzübergängen intern zu intern im Einsatz?

- Ja
- Nein

7.4 Haben Sie NIDS an den Netzübergängen Ja Nein intern zu extern im Einsatz?

- Ja
- Nein

7.5 Haben Sie NIDS innerhalb von internen Netzbereichen im Einsatz?

- Ja
- Nein

7.6 In welcher Form werden die Netzwerkdaten in Ihrer Organisation analysiert?

- Auswertung des gesamten Netzwerkverkehrs
- Metadaten
- Firewall-Daten
- Netflow-Daten
- Andere (Freitext)

7.7 Welche?
(Freitext)

7.8 Bitte benennen Sie konkret Produkte / Anwendungen, die bei Ihnen zur Angriffserkennung zum Einsatz kommen.
(Freitext)

7.9 Kommt bei Ihnen ein Dienstleister (Managed Security Service Provider) zur Angriffserkennung zum Einsatz?
Nein, und auch nicht vorgesehen
Nein, aber in Prüfung
Nein, aber in Vorbereitung / Initialisierung
Ja, bereits produktiv für das Monitoring
Ja, aber soll wieder eingestellt werden

7.10 Welcher Dienstleister kommt bei Ihnen als MSSP zum Einsatz (bitte Unternehmen konkret benennen)?
(Freitext)

7.11 Welche Anteile (welche Teilprozesse) eines SOC sind / werden an einen MSSP ausgelagert (S. 14ff des MITRE- Dokumentes gibt weiteren Input zu den Teilprozessen)?

- Incident Triage, Analysis, and Response
- Vulnerability Management SOC Management
- Cyber Threat Intelligence, Hunting, and Analytics
- SOC Tools, Architecture, and Engineering
- Pen-Testing / Red-Teaming
- Situational Awareness, Communications, and Training

- Andere (bitte spezifizieren)

7.12 Welche?

(Freitext)

7.13 Welche Bereiche werden von Ihren SzA heute bereits abgedeckt?

- IT (komplett)
- IT (angefangen)
- MT (komplett)
- MT (angefangen)
- VT (komplett)
- VT (angefangen)

7.14 Welche Gefährdungen adressieren Sie aktuell mit Ihrer SzA-Umsetzung? Bei einer hohen Anzahl konzentrieren Sie sich bitte auf die Top-5.

(Freitext)

8. Branchenspezifische Regeln und Normen

8.1 Welche Auflagen (Gesetze, Verordnungen & Co) werden über das ISMS bedient?

- BSIG
- Deutsches Medizinproduktegesetz (MPG)
- SGB
- Medizinprodukte- Sicherheitsplanverordnung (MPSV)
- EU-Medizinprodukteverordnung (Medical Device Regulation - MDR)
- Medizinprodukte- Betreiberverordnung (MPBetreibV)
- Andere (bitte spezifizieren)

8.2 Welche?

(Freitext)

8.3 Welche Standards kommen zur Umsetzung ISMS und IT-Sicherheit zur Anwendung?

- ISO27000
- IT-Grundschutz / BSI 200-x
- B3S
- Andere (bitte spezifizieren)

8.4 Welche?

(Freitext)

8.5 Welche Good Practice stellen konkret Anforderungen an SzA bzw. liefern Input zur Ausgestaltung und werden herangezogen?

- BSI: OH SzA
- BSI: Mindeststandard zur Protokollierung und Detektion von Cyber-Angriffen
- BSI: Monitoring und Anomalieerkennung in Produktionsnetzwerken
- NIST: Good Practice on Computer Security Incident Handling
- ENISA: How to set up CSIRT and SOC
- MITRE: 11 Strategies of a World-Class Cybersecurity Operations Center
- Andere (bitte spezifizieren)

8.6 Welche?
(Freitext)

9. Anonymisierung / Pseudonymisierung von Protokoll(ierungs)daten und Datenschutz

9.1 Ist das SIEM in das Verzeichnis der Ja Nein Verfahrenstätigkeiten aufgenommen?

- Ja
- Nein

9.2 Wie lange werden die Rohdaten der Protokollierung aufbewahrt? Bitte geben Sie die Einheit (Tage, Woche, Monate) mit an.
(Freitext)

9.3 Wie lange werden die bearbeiteten Protokollierungsdaten (qualifizierte Ereignisse) aufbewahrt? Bitte geben Sie die Einheit (Tage, Woche, Monate) mit an.
(Freitext)

9.4 Werden Protokolldaten (Datenerhebung in Ja Nein Netzverkehr) auch auf TCP/IP-Layer 7 in das Monitoring einbezogen?

- Ja
- Nein

9.5 Sofern Patientendaten von der Protokollierung erfassts sind, können Sie kurz den Umgang damit beschreiben?
(Freitext)

9.6 Wie viele Personen haben Zugriff auf die zentralen Protokollierungsdaten?
(Freitext)

9.7 Werden Zugriffe auf das SIEM konsequent festgehalten und gibt es hierfür Auswertemöglichkeiten (Anwendungsprotokollierung des SIEM)?

- Ja, inkl. Auswertemöglichkeiten
- Ja, werden festgehalten
- Nein
- Unbekannt

10. Protokollierungsfähigkeit von Medizintechnik

Nachfolgende Fragen beziehen sich ausschließlich auf die Umsetzung der Angriffserkennung im Netz der netzwerkfähigen Medizintechnik.

10.1 Gemessen an der Gesamtzahl der in die Protokollierung einzubeziehenden Systeme: wieviel Systeme können Systemprotokollierung (Syslog oder Alternative) versenden?

- <10%
- 10% - 25%
- 25% - 50%
- 50% - 75%
- 75% - 90%
- >90%
- Einschätzung nicht möglich

10.2 Wie viele Systeme liefern konkret Systemprotokollierung in das zentrale Monitoring ein? Das Quellsystem kann hierbei auf verschiedenen Wegen (bspw. Remote-Syslog oder Agent der SIEM-Lösung) eingebunden sein.
(Freitext)

10.3 Fuhren Sie ein Netzmonitoring auf Basis von Netflow durch?

- Ja
- Nein

10.4 Fuhren Sie ein Netzmonitoring mit Hilfe von Ja Nein NIDS/NIPS durch?

- Ja
- Nein

10.5 Fuhren Sie ein Netzmonitoring mit Hilfe von PCAP-Mitschnitten durch?

- Ja
- Ja, aber nur anlassbezogen

- Nein

10.6 Achten Sie bei der Beschaffung von netzwerkfähiger Medizintechnik darauf, ob Sicherheitseigenschaften und konkret auch Systemprotokollierung (Syslog & co) unterstützt wird?

- Nein
- Ja, Sicherheitseigenschaften
- Ja, Sicherheitseigenschaften inkl. Systemprotokollierung

10.7 Ziehen Sie zur Bewertung von Sicherheitseigenschaften bei der Beschaffung von Medizintechnik Herstelleraussagen im MDS2-Format heran?

- Ja
- Nein

11. Automatisierungspotentiale der Reaktion

11.1 Wie viele Playbooks bzw. Runbooks (Beschreibungen zur standardisierten Reaktion) sind in Ihrer Sicherheitsorganisation verschriftlicht?

- 0
- <5
- 6 – 20
- 21 – 50
- >50

11.2 Findet in Ihrer Organisation eine automatisierte Ja Nein Reaktion statt?

- Ja
- Nein

11.3 Wenn ja, automatisierte Reaktion in der IT?

- Ja
- Nein

11.4 Wenn ja, automatisierte Reaktion in der VT?

- Ja
- Nein

11.5 Wenn ja, automatisierte Reaktion in der MCPS?

- Ja
- Nein

11.6 Kommt in Ihrer Organisation Security Orchestration Automation and Responses (SOAR) zum Einsatz?

- Ja
- Nein

11.7 Welches Produkt?
(Freitext)

12.1 Die OH überlässt dem KRITIS-Betreiber, ob die Erkennung von Ereignissen manuell oder automatisiert erfolgt. Welchen Anteil hat bei Ihnen die automatische Erkennung

- Erkennung findet nur automatisiert statt (manuelle Tätigkeit nur bei Alarm)
- Erkennung primär automatisiert; in Einzelfällen auch manuelles Screening von Protokollierungsdaten
- Automatisierte und manuelle Auswertung halten sich in etwa in Waage
- Erkennung primär manuell, erste Regelsätze unterstützen bei der Automatisierung der Erkennung
- Keine Automatisierung - Sichtung von Protokollierungsdaten erfolgt nur manuell

12.2 Laut OH müssen Reaktionszeiten für Alarme der SzA definiert sein, welche sich wiederum aus der Risikobewertung speisen müssen. Welche Reaktionszeiten sind bei Ihnen definiert?

- Innerhalb von Minuten (24 x 7)
- Innerhalb von Minuten (8 x 5)
- Innerhalb weniger Stunden
- Innerhalb eines Arbeitstages
- Mehr als ein Arbeitstag
- Noch nicht definiert

12.3 Mitarbeitende für SzA müssen namentlich benannt sein. Wie viele Mitarbeiter sind in Ihrer Organisation für das Thema SzA benannt?
(Freitext)

12.4 Welchem organisatorischen Arbeitsbereich ist CISO IT / CIO andere (bitte die Angriffserkennung mit den Mitarbeitenden / spezifizieren) Dienstleistern zugeordnet?

- CISO
- IT/CIO
- Andere (bitte spezifizieren)

12.5 Welcher?
(Freitext)

12.6 Welche Angebote / Anbieter (z.B. von Schulungen) nutz(t)en Sie, um die Mitarbeitenden für die Tätigkeiten in der Angriffserkennung vorzubereiten?
(Freitext)

12.7 Haben Sie noch weitere Anmerkungen oder Hinweise, die für die Ausgestaltung des B3S hinsichtlich der Thematik "Systeme zur Angriffserkennung" relevant sind und aus Ihrer Sicht in dieser Befragung zu kurz gekommen sind?
(Freitext)

Anlage 3: Interviewleitfaden Expertengespräche

Themenfeld 1: Validierung von SdT-Empfehlungen
Historische Festlegung des Standes der Technik (SdT):
Auf welcher Grundlage und nach welchen Kriterien wurde der Stand der Technik (durch den AK SdT) in der Vergangenheit bei anderen IT-Themen bestimmt?
Wurden (internationale) Normen / Good practices herangezogen? Welche (primär)? Inwieweit spielt der IT-Grundschutz eine Rolle?
Mit welchen Verfahren und Methoden wurde validiert, ob Empfehlungen tatsächlich dem SdT entsprechen?
Einsicht in Entscheidungsprozesse beim BSI:
Hat das Bundesamt für Sicherheit in der Informationstechnik Einblicke in seine Entscheidungsprozesse gegeben - insbesondere darüber, wie beurteilt wurde, ob Teile des B3S dem SdT entsprechen?
Gab es Fälle, in denen der SdT für Branchentechnik eingeschränkt wurde? Wie wurde dies bewertet?
Einfluss neuer Technologien:
Welche Rolle spielen neuartige, potenziell disruptive Technologien, die von Sicherheitsfirmen bereits angeboten, aber noch nicht von Regulierungsbehörden geprüft und in Empfehlungen aufgenommen wurden? (Randnotiz: z.B.: Asimily oder Cybersense)

Themenfeld 2: Anwendbarkeit von SdT-Empfehlungen
Bewertung des SdT in Bezug auf Unterschiede zwischen deutschen Krankenhäusern
War bei der Erstellung der ersten Version des B3S absehbar, dass kleinere Krankenhäuser durch die Regelungen des § 75c SGB V für den SdT ebenfalls auf den B3S verwiesen werden? (Randnotiz: Nach Timeline eher nicht - Feststellungsbescheid B3S 1.0 22.10.2019; Verpflichtung nach 75c des SGB V ab 01.01.22 (verkündet am 20.10.2020)
Falls ja: Wie wurde der Stand der Technik unter Berücksichtigung der deutlichen Unterschiede zwischen deutschen Krankenhäusern abgewogen?
Feedback und Herausforderungen bei der Umsetzung des B3S:
Welche Rückmeldungen oder Hinweise gab es von Krankenhäusern bezüglich der Umsetzung der Sicherheitsanforderungen aus dem B3S v1.0, insbesondere von solchen, die erst durch die Einführung des § 75c SGB V darauf aufmerksam wurden?
Wie bewerten diese Krankenhäuser die Nützlichkeit des B3S? Welche Kernprobleme wurden identifiziert und wie wurde darauf im Arbeitskreis SdT reagiert?
Sollten in einer zukünftigen Version unterschiedliche Maßstäbe für kleinere und größere Häuser angewendet werden, oder sollten Empfehlungen möglichst für alle Häuser gelten?
Falls letzteres: „Welche Elemente des B3S sollten flexibel gestaltet werden, um unterschiedlichen Anforderungen gerecht zu werden?

Themenfeld 3: Erfahrungswerte bei der Umsetzung von IT-Sicherheit
Erfahrungswerte aus der Umsetzung bei SzA:
Welche Erfahrungen und Erkenntnisse wurden bei der Implementierung von SzA bisher gesammelt?
Erkenntnisse aus der Umsetzung anderer IT-Sicherheitsthemen im Krankenhaus:
Welche Herausforderungen traten bei der Implementierung von IT-Sicherheitsmaßnahmen auf und wie wurden diese bewältigt? Welche Lösungen oder Workarounds wurden entwickelt?
Inwiefern beeinflussen IT-Sicherheitsmaßnahmen das Tagesgeschäft (sowohl positiv als auch negativ)?
Hat die Inanspruchnahme externer Unterstützung beim Aufbau und Betrieb von Sicherheitsmaßnahmen zu erkennbaren Verbesserungen geführt?
Lessons Learned und Best Practices:
Wurden 'Lessons Learned' oder 'Best Practices' identifiziert, die sich aus den bisherigen Umsetzungen von IT-Sicherheitsmaßnahmen ergeben haben?

Chapter 8

Conclusion

This dissertation has advanced the understanding and applicability of attack detection within hospital environments, addressing the critical need for tailored and effective security measures in an industry, where patient safety and operational continuity are paramount. It systematically outlines legal requirements and dependencies, providing a clear rationale for the necessity of implementing SzA in hospitals. The current legislative framework empowers the specification of *state of the art* requirements adapted to specific industries, recognizing their unique operational challenges. By grounding this specification for hospital attack detection in a legally sound and methodically transparent approach, this work provides a clear pathway to meaningful security improvements. Central to this dissertation is the recognition that current generic standards and guidelines, while foundational, must be contextualized for hospitals' unique operational environments. A survey of German hospitals and subsequent expert interviews highlighted the practical challenges impeding the direct adoption of general SzA principles.

In parallel, the *state of research and technology* of attack detection for MCPS was assessed in an extensive literature review. The findings reveal that certain industry-specific characteristics are already being addressed in current research. For instance, the heightened focus on insider threats in the context of hospitals recognizes the elevated risk posed by patients and visitors – an issue less pronounced in other KRITIS sectors. Similarly, most detection efforts concentrate on network-level mechanisms, which reflects an awareness of the unique challenges posed by MT and VT, as the devices in these areas frequently do not have the necessary capabilities for on-device detection. For future research, the major necessity is publicly available research datasets to facilitate the development and testing of new attack detection approaches. There is also great potential in the use of healthcare data for attack detection. Current advancements in this domain are still too immature for practical implementation in hospitals, requiring further refinement before they can contribute to operational security strategies. As this dissertation exclusively reviews the *state of research and technology* in the field of MCPS, further assessments of the *state of research and technology*, particularly in VT and specialized hospital IT systems, would help to identify additional emerging trends. Insights from such studies could introduce new measures to advance attack detection in the healthcare sector.

As an additional component, the review of existing standards and best practices informed the development of short- and mid-term measures for hospital attack detection. These transitional measures serve as interim solutions, fostering protection until more sustainable and comprehen-

sive industry solutions are established. They enable effective attack detection without relying on immediate modifications by MT and VT manufacturers or adjustments to certification processes. Achieving long-term progress will necessitate enhancing logging and data collection capabilities in MT and VT devices to provide sufficient data for identifying suspicious activities. Building on this, analysis mechanisms must be developed that utilize these logs and address industry-specific threats and advanced attack vectors to secure MCPS comprehensively. Until such advancements are realized, the transitional measures are needed.

As part of the proposed short-term measures, this dissertation provides actionable strategies to assist hospitals in securing their networks. One particularly promising approach involves leveraging MDS2 documents, which provide detailed security-related properties of medical devices. Leveraging this information to streamline device integration into detection infrastructures can alleviate operational burdens and improve situational awareness. Aggregating and correlating MDS2 data enables evidence-based decision-making, reveals device-spanning deficiencies, and addresses common misconceptions and incorrect assumptions affecting both operators and researchers. This contributes to a better understanding of device-specific and systemic vulnerabilities, offering actionable insights or immediate improvement in hospital security. Regarding future research, it would be beneficial to further develop the MDS2 standard and create a centralized repository for storing, retrieving, and analyzing MDS2 documents. Such a centralized MDS2 repository would serve as a valuable resource for researchers and operators, enabling them to access accurate, real-world data about the security properties of medical devices. This would facilitate the identification of new and widespread problems while ensuring that emerging research approaches are based on accurate and relevant assumptions, driving more effective and practical advancements in the field.

Honeypots emerge as a promising mid-term strategy, able to enhance security in hospitals without relying on manufacturer capabilities or adjustments to certification processes. They can be implemented independently of existing security properties, directly improving protection in their network segments. While their immediate, widespread deployment remains infeasible in specialized domains due to the complexity of creating honeypots that accurately mimic devices within these domains, this dissertation explores the potential of using LLMs as honeypot backends. By investigating their applicability, this research contributes to a more efficient path for developing realistic and adaptable honeypots. A unified evaluation approach for LLMs is given, enabling researchers to compare the performance of newly developed or fine-tuned models. This work lays the foundation for integrating LLMs into security frameworks. Despite the encouraging findings, substantial challenges remain before LLM-based honeypots can be deployed in production. GPT-3.5, as a representative LLM, struggles to maintain context during SSH sessions, a critical factor for convincingly simulating real systems. Furthermore, potential vulnerabilities, such as prompt injection attacks, require further research. Additionally, the applicability of LLMs for simulating medical-specific protocols and devices has yet to be thoroughly tested. Nevertheless, with continued research and development, LLM-based honeypots could emerge as a viable mid-term solution, offering healthcare networks an adaptable and realistic approach to deception-based security. Their potential to enhance threat detection

and mitigate risks makes them a promising avenue for securing hospital systems in the coming years.

By bridging the gap between regulatory mandates, research insights, and practical hospital operations, this dissertation sets a foundation for continuous improvement for SzA. The results of this dissertation are now being used to revise the industry-specific security standard for hospitals in Germany. The transitional short- and mid-term solutions can evolve into long-term strategies, ensuring that hospital security measures keep pace with the rapid technological innovations and emerging threats that define the modern healthcare landscape. In doing so, this work contributes to a more secure and resilient healthcare infrastructure sustaining public trust. Preventing incidents like the 2020 attack on the UKD is vital to safeguard patient lives in the future.

Bibliography

Abbas-Escribano, Marwan and Hervé Debar (2023). "An improved honeypot model for attack detection and analysis". In: *Proceedings of the 18th International Conference on Availability, Reliability and Security*, pp. 1–10 (cit. on p. 23).

Almaiah, Mohammed Amin, Aitizaz Ali, Fahima Hajjej, Muhammad Fermi Pasha, and Manal Abdullah Alohali (Mar. 2022). "A Lightweight Hybrid Deep Learning Privacy Preserving Model for FC-Based Industrial Internet of Medical Things". en. In: *Sensors* 22.6, p. 2112. ISSN: 1424-8220. DOI: 10.3390/s22062112. URL: https://www.mdpi.com/1424-8220/22/6/2112 (visited on 04/21/2022) (cit. on p. 19).

Anderson, James P (1980). "Computer security threat monitoring and surveillance". In: *Technical Report, James P. Anderson Company* (cit. on p. 5).

Astillo, Philip Virgil, Daniel Gerbi Duguma, Hoonyong Park, Jiyoon Kim, Bonam Kim, and Ilsun You (Mar. 2022). "Federated intelligence of anomaly detection agent in IoTMD-enabled Diabetes Management Control System". en. In: *Future Generation Computer Systems* 128, pp. 395–405. ISSN: 0167-739X. DOI: 10.1016/j.future.2021.10.023. URL: https://www.sciencedirect.com/science/article/pii/S0167739X21004192 (visited on 04/21/2022) (cit. on p. 19).

Bhatt, Sandeep, Pratyusa K Manadhata, and Loai Zomlot (2014). "The operational role of security information and event management systems". In: *IEEE security & Privacy* 12.5, pp. 35–41 (cit. on p. 8).

Biswas, Som (Feb. 2023). "Role of ChatGPT in Computer Programming." In: 2023, pp. 9–15. DOI: 10.58496/MJCSC/2023/002. URL: https://mesopotamian.press/journals/index.php/cs/article/view/51 (cit. on p. 24).

Bundesamt für Sicherheit in der Informationstechnik (Sept. 26, 2022). *Orientierungshilfe zum Einsatz von Systemen zur Angriffserkennung*. Technical Report. Version 1.0. Accessed: October 15, 2024. Bundesamt für Sicherheit in der Informationstechnik. URL: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/KRITIS/oh-sza.pdf (cit. on p. 10).

– (2024a). *Industry-specific security standards (B3S)*. Accessed: 2024-11-28. URL: https://www.bsi.bund.de/EN/Themen/Regulierte-Wirtschaft/Kritische-Infrastrukturen/Allgemeine-Infos-zu-KRITIS/Stand-der-Technik-umsetzen/Branchenspezifische-Sicherheitsstandards-B3S/branchenspezifische-sicherheitsstandards-b3s_node.html (cit. on p. 11).

Bundesamt für Sicherheit in der Informationstechnik (2024b). *UP KRITIS*. Accessed: 28 November 2024. URL: https://www.bsi.bund.de/EN/Themen/Regulierte-Wirtschaft/Kritische-Infrastrukturen/UP-KRITIS/up-kritis_node.html (cit. on p. 11).

– (2024c). *UP KRITIS*. Accessed: 28 November 2024. URL: https://www.bsi.bund.de/DE/Themen/Regulierte-Wirtschaft/Kritische-Infrastrukturen/UP-KRITIS/Branchenarbeitskreise/branchenarbeitskreise_node.html (cit. on p. 11).

Bundesministerium der Justiz (2024). *Handbuch der Rechtsförmlichkeit*. German. Accessed: Oktober 26, 2024. Bundesministerium der Justiz. URL: https://www.bmj.de/SharedDocs/Publikationen/DE/Fachpublikationen/Handbuch_der_Rechtsfoermlichkeit.html (cit. on pp. 15, 30).

Bundestag, Deutscher (2009). *Gesetz über das Bundesamt für Sicherheit in der Informationstechnik (BSI-Gesetz - BSIG)*. BGBl. I S. 2821. Amended by Article 12 of the law from June 23, 2021 (BGBl. I S. 1982) (cit. on p. 11).

– (2020). *Gesetz zum Schutz elektronischer Patientendaten in der Telematikinfrastruktur*. https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/Gesetze_und_Verordnungen/GuV/P/PDSG_bgbl.pdf (cit. on p. 11).

– (2021). *Zweites Gesetz zur Erhöhung der Sicherheit informationstechnischer Systeme (IT-Sicherheitsgesetz 2.0)* (cit. on pp. 2, 9, 15).

– (Mar. 25, 2024). *Gesetz zur Beschleunigung der Digitalisierung des Gesundheitswesens (Digital-Gesetz — DigiG)*. https://www.recht.bund.de/bgbl/1/2024/101/VO.html. Gesetz. BGBl. 2024 I Nr. 101 vom 25.03.2024 (cit. on p. 11).

Bundestag, Deutscher (2024). *Entwurf eines Gesetzes zur Umsetzung der NIS-2-Richtlinie und zur Regelung wesentlicher Grundzüge des Informationssicherheitsmanagements in der Bundesverwaltung (NIS-2-Umsetzungs- und Cybersicherheitsstärkungsgesetz)*. Technical Report: 20/13184. Deutscher Bundestag. URL: https://dserver.bundestag.de/btd/20/131/2013184.pdf (visited on 11/26/2024) (cit. on p. 10).

Check Point Research (2023). *38% Increase in 2022 Global Cyberattackss*. Check Point Software Technologies Ltd. URL: https://blog.checkpoint.com/2023/01/05/38-increase-in-2022-global-cyberattacks/ (visited on 11/26/2024) (cit. on p. 1).

Clark, Shane S. and Kevin Fu (2012). "Recent Results in Computer Security for Medical Devices". en. In: *Wireless Mobile Communication and Healthcare*. Ed. by Konstantina S. Nikita, James C. Lin, Dimitrios I. Fotiadis, and Maria-Teresa Arredondo Waldmeyer. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Berlin, Heidelberg: Springer, pp. 111–118. ISBN: 978-3-642-29734-2. DOI: 10.1007/978-3-642-29734-2_16 (cit. on p. 20).

Cohen, Fred (1999). *The Deception ToolKit (DTK)*. Online. Accessed: 2024-10-16. URL: http://all.net/dtk/ (cit. on p. 22).

CrowdStrike (2023). *2023 Global Threat Report*. Technical Report. CrowdStrike. URL: https://www.crowdstrike.com/wp-content/uploads/2023/02/2023-Global-Threat-Report-Executive-Summary.pdf (cit. on p. 1).

Decker, E, R Wood, S Mohiuddin, D Nock, and A Venugopalan (2023). *Hospital Cyber Resiliency Initiative Landscape Analysis. HHS 405 (d)*. Technical Report: URL: https://405d.hhs.gov/Documents/405d-hospital-resiliency-analysis.pdf (cit. on pp. 1, 2).

Denning, Dorothy E (1987). "An intrusion-detection model". In: *IEEE Transactions on software engineering* 2, pp. 222–232 (cit. on p. 5).

Deutsche Krankenhausgesellschaft (Dec. 2022). *Branchenspezifischer Sicherheitsstandard für die medizinische Versorgung in Krankenhäusern*. Technical Report: version 1.2. Eignung nach § 8a Abs. 1 und Abs. 1a BSIG festgestellt am 10.1.2023. Deutsche Krankenhausgesellschaft. URL: https://www.dkgev.de/themen/digitalisierung-daten/informationssicherheit-und-technischer-datenschutz/informationssicherheit-im-krankenhaus/ (cit. on pp. 11–14).

ENISA (July 2023). *ENISA Threat Landscape: Health Sector*. DOI: 10.2824/163953. URL: https://www.enisa.europa.eu/publications/enisa-threat-landscape-health-sector (cit. on pp. 1, 2).

Fernández Maimó, Lorenzo, Alberto Huertas Celdrán, Angel L. Perales Gómez, Félix J. García Clemente, James Weimer, and Insup Lee (Jan. 2019). "Intelligent and Dynamic Ransomware Spread Detection and Mitigation in Integrated Clinical Environments". en. In: *Sensors* 19.5. Number: 5 Publisher: Multidisciplinary Digital Publishing Institute, p. 1114. ISSN: 1424-8220. DOI: 10.3390/s19051114. URL: https://www.mdpi.com/1424-8220/19/5/1114 (visited on 04/21/2022) (cit. on p. 20).

Forescout (2024a). *The Riskiest Connected Devices in 2024*. Accessed: 2024-11-13. URL: https://www.forescout.com/resources/2024-riskiest-connected-devices/ (cit. on p. 24).

– (2024b). *Unveiling the Persistent Risks of Connected Medical Devices*. Accessed: 2024-11-13. URL: https://www.forescout.com/resources/iomt-persistent-risk-report/ (cit. on p. 23).

Franco, Javier, Ahmet Aris, Berk Canberk, and A. Selcuk Uluagac (2021). "A Survey of Honeypots and Honeynets for Internet of Things, Industrial Internet of Things, and Cyber-Physical Systems". In: *IEEE Communications Surveys & Tutorials* 23.4, pp. 2351–2383. DOI: 10.1109/COMST.2021.3106669 (cit. on pp. 22, 23).

Fuchsberger, Andreas (2005). "Intrusion detection systems and intrusion prevention systems". In: *Information Security Technical Report* 10.3, pp. 134–139 (cit. on p. 7).

George, Dr A Shaji, AS Hovan George, T Baskar, and Digvijay Pandey (2021). "XDR: The Evolution of Endpoint Security Solutions-Superior Extensibility and Analytics to Satisfy the Organizational Needs of the Future". In: *International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)* 8.1, pp. 493–501 (cit. on p. 8).

Gupta, Deepti, Olumide Kayode, Smriti Bhatt, Maanak Gupta, and Ali Saman Tosun (Nov. 2021). "Hierarchical Federated Learning based Anomaly Detection using Digital Twins for Smart Healthcare". en. In: *arXiv:2111.12241 [cs]*. arXiv: 2111.12241. URL: http://arxiv.org/abs/2111.12241 (visited on 04/22/2022) (cit. on p. 19).

Hady, Anar A., Ali Ghubaish, Tara Salman, Devrim Unal, and Raj Jain (2020). "Intrusion Detection System for Healthcare Systems Using Medical and Network Data: A Comparison Study". en. In: *IEEE Access* 8, pp. 106576–106584. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.3000421. URL: https://ieeexplore.ieee.org/document/9109651/ (visited on 04/22/2022) (cit. on p. 20).

Hameed, Shilan S., Wan Haslina Hassan, and Liza Abdul Latiff (2021). "An Efficient Fog-Based Attack Detection Using Ensemble of MOA-WMA for Internet of Medical Things". en. In: *Innovative Systems for Intelligent Health Informatics*. Ed. by Faisal Saeed, Fathey Mohammed, and Abdulaziz Al-Nahari. Lecture Notes on Data Engineering and Communications Technologies. Cham: Springer International Publishing, pp. 774–785. ISBN: 978-3-030-70713-2. DOI: 10.1007/978-3-030-70713-2_70 (cit. on p. 19).

Health Level Seven International (May 2024). *2024 State of FHIR Survey Results*. https://www.hl7.org/documentcenter/public/white-papers/2024%20StateofFHIRSurveyResults_final.pdf. Accessed: 2024-11-17 (cit. on p. 24).

Health Sector Cybersecurity Coordination Center (Apr. 2024). *HC3's Top 10 Most Active Ransomware Groups*. Analyst Note 202404051700. Accessed Oktober 24, 2024. U.S. Department of Health and Human Services. URL: https://www.aha.org/system/files/media/file/2024/04/HC3-tlp-clear-analyst-note-top-10-most-active-ransomware-groups-4-10-24.pdf (cit. on p. 1).

Hu, Yuqi, Siyu Cheng, Yuanyi Ma, Shuangwu Chen, Fengrui Xiao, and Quan Zheng (2024). "MySQL-Pot: A LLM-Based Honeypot for MySQL Threat Protection". In: *2024 9th International Conference on Big Data Analytics (ICBDA)*. IEEE, pp. 227–232 (cit. on p. 25).

Joyia, Gulraiz, M. Akram, Chaudary Akbar, and Muhammad Maqsood (Jan. 2018). "Evolution of Health Level-7: A Survey". In: pp. 118–123. DOI: 10.1145/3178461.3178480 (cit. on p. 24).

Justiz, Bundesamt für (2016). *Verordnung zur Bestimmung Kritischer Infrastrukturen nach dem BSI-Gesetz (BSI-Kritisverordnung - BSI-KritisV)*. vom 22. April 2016 (BGBl. I S. 958), die zuletzt durch Artikel 1 der Verordnung vom 29. November 2023 (BGBl. 2023 I Nr. 339) geändert worden ist. URL: https://www.gesetze-im-internet.de/bsi-kritisv/ (cit. on pp. 2, 10).

Karantzas, George and Constantinos Patsakis (2021). "An empirical assessment of endpoint detection and response systems against advanced persistent threats attack vectors". In: *Journal of Cybersecurity and Privacy* 1.3, pp. 387–421 (cit. on p. 8).

Keri, Mikael (2023). *Dicompot - A Digital Imaging and Communications in Medicine (DICOM) Honeypot*. `https://github.com/nsmfoo/dicompot`. Accessed: 2024-10-03 (cit. on p. 23).

Khraisat, Ansam, Iqbal Gondal, Peter Vamplew, and Joarder Kamruzzaman (2019). "Survey of intrusion detection systems: techniques, datasets and challenges". In: *Cybersecurity* 2.1, pp. 1–22 (cit. on p. 6).

Kitchenham, Barbara, Stuart Charters, et al. (2007). "Guidelines for performing systematic literature reviews in software engineering version 2.3". In: *Engineering* 45.4ve, p. 1051 (cit. on pp. 29, 30).

Knerler, Kathryn, Ingrid Parker, and Carson Zimmerman (2023). *11 Strategies of a World-Class Cybersecurity Operations Center*. MITRE (cit. on pp. 7, 9, 21).

Lunt, Teresa F, Ann Tamaru, Fred Gilham, R Jagannathan, Caveh Jalali, Peter G Neumann, Harold S Javitz, Alfonso Valdes, and Thomas D Garvey (1992). *A real-time intrusion-detection expert system (IDES)*. Citeseer (cit. on p. 5).

McKee, Forrest and David Noever (2023). "Chatbots in a honeypot world". In: *arXiv preprint arXiv:2301.03771* (cit. on p. 25).

Meng, Weizhi, Wenjuan Li, and Liqiu Zhu (Nov. 2020). "Enhancing Medical Smartphone Networks via Blockchain-Based Trust Management Against Insider Attacks". In: *IEEE Transactions on Engineering Management* 67.4. Conference Name: IEEE Transactions on Engineering Management, pp. 1377–1386. ISSN: 1558-0040. DOI: 10.1109/TEM.2019.2921736 (cit. on p. 19).

Mfogo, Volviane Saphir, Alain Zemkoho, Laurent Njilla, Marcellin Nkenlifack, and Charles Kamhoua (2023). "AIIPot: Adaptive intelligent-interaction honeypot for IoT devices". In: *2023 IEEE 34th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. IEEE, pp. 1–6 (cit. on p. 25).

Miliard, Mike (Sept. 2020). *Hospital ransomware attack leads to fatality after causing delay in care*. Accessed November 20, 2024. URL: `https://www.healthcareitnews.com/news/hospital-ransomware-attack-leads-fatality-after-causing-delay-care` (cit. on p. 1).

Mitchell, Robert and Ing-Ray Chen (Jan. 2015). "Behavior Rule Specification-Based Intrusion Detection for Safety Critical Medical Cyber Physical Systems". In: *IEEE Transactions on Dependable and Secure Computing* 12.1. Conference Name: IEEE Transactions on Dependable and Secure Computing, pp. 16–30. ISSN: 1941-0018. DOI: 10.1109/TDSC.2014.2312327 (cit. on p. 19).

Mukherjee, Biswanath, L Todd Heberlein, and Karl N Levitt (1994). "Network intrusion detection". In: *IEEE network* 8.3, pp. 26–41 (cit. on p. 6).

Newaz, AKM Iqtidar, Amit Kumar Sikder, Mohammad Ashiqur Rahman, and A. Selcuk Uluagac (Oct. 2019). "HealthGuard: A Machine Learning-Based Security Framework for Smart Healthcare Systems". In: *2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, pp. 389–396. DOI: 10.1109/SNAMS.2019.8931716 (cit. on p. 19).

Ponemon (Jan. 2023). *The Impact of Ransomware on Patient Safety and the Value of Cybersecurity Benchmarking*. Technical Report: Ponemon Institute. URL: https://dd80b675424c132b90b3-e48385e382d2e5d17821a5e1d8e4c86b.ssl.cf1.rackcdn.com/external/ponemon-report-jan-2023.pdf (cit. on p. 2).

Ralston, William (Nov. 2020). "The untold story of a cyberattack, a hospital and a dying woman". In: *Wired*. Accessed November 20, 2024. URL: https://www.wired.com/story/ransomware-hospital-death-germany/ (cit. on p. 1).

Ray, Partha Pratim (2023). "ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope". In: *Internet of Things and Cyber-Physical Systems* 3, pp. 121–154. ISSN: 2667-3452. DOI: https://doi.org/10.1016/j.iotcps.2023.04.003. URL: https://www.sciencedirect.com/science/article/pii/S266734522300024X (cit. on p. 24).

Schmall, Markus (2022). *medpot: Medical Protocol Honeypot*. https://github.com/schmalle/medpot. Accessed: 2024-10-03 (cit. on p. 23).

Seo, Gihong, Sewon Park, and Munjae Lee (Sept. 14, 2022). "How to calculate the life cycle of high-risk medical devices for patient safety". In: *Frontiers in Public Health* 10, p. 989320. ISSN: 2296-2565. DOI: 10.3389/fpubh.2022.989320. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9515981/ (visited on 08/17/2024) (cit. on p. 13).

Snapp, Steven R, James Brentano, Gihan Dias, Terrance L Goan, L Todd Heberlein, Che-Lin Ho, and Karl N Levitt (1991). "DIDS (distributed intrusion detection system)-motivation, architecture, and an early prototype". In: (cit. on p. 7).

Sophos (2024). *The State of Ransomware in Healthcare 2024*. Whitepaper. Accessed November 30, 2024. Sophos. URL: https://www.sophos.com/en-us/whitepaper/state-of-ransomware-in-healthcare (cit. on p. 2).

Stein, Stefan, Simon Weber, Michael Pilgermann, Thomas Schrader, and Martin Sedlmayr (2024). "A Novel Approach to Medical Device IT Security Landscape Analysis Leveraging Manufacturer Disclosure Statements". In: *IEEE Access* (cit. on p. 67).

Stoll, C. (1989). *The Cuckoo's Egg: Tracking a Spy Through the Maze of Computer Espionage*. Doubleday. ISBN: 9780385249461 (cit. on p. 22).

Thamilarasu, Geethapriya, Adedayo Odesile, and Andrew Hoang (2020). "An Intrusion Detection System for Internet of Medical Things". In: *IEEE Access* 8. Conference Name: IEEE Access, pp. 181560–181576. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.3026260 (cit. on p. 19).

U.S. Department of Health and Human Services (2023). *Health Industry Cybersecurity Practices: Managing Threats and Protecting Patients.* Developed as part of the 405(d) Task Group initiative to provide cybersecurity guidelines for the healthcare sector. URL: https://405d.hhs.gov/Documents/HICP-Main-508.pdf (cit. on p. 16).

Vasilatos, Christoforos, Dunia J Mahboobeh, Hithem Lamri, Manaar Alam, and Michail Maniatakos (2024). "LLMPot: Automated LLM-based Industrial Protocol and Physical Process Emulation for ICS Honeypots". In: *arXiv preprint arXiv:2405.05999* (cit. on p. 25).

Wagner, Thomas D, Khaled Mahbub, Esther Palomar, and Ali E Abdallah (2019). "Cyber threat intelligence sharing: Survey and research directions". In: *Computers & Security* 87, p. 101589 (cit. on p. 9).

Weber, Simon, Michael Pilgermann, Stefan Stein, and Thomas Schrader (2024a). "SzA4Hosp-Systeme zur Angriffserkennung in der Medizinischen Versorgung". In: (cit. on pp. 12, 13, 81).

Weber, Simon B, Marc Feger, and Michael Pilgermann (2024b). "Don't Stop Believin': A Unified Evaluation Approach for LLM Honeypots". In: *IEEE Access* (cit. on pp. 25, 55).

Weber, Simon B, Stefan Stein, Michael Pilgermann, and Thomas Schrader (2023). "Attack detection for medical cyber-physical systems–a systematic literature review". In: *IEEE Access* 11, pp. 41796–41815 (cit. on pp. 13, 29).

WHO and Statista (Nov. 2024). *Number of attacks reported annually against healthcare worldwide from 2015 to 2024 [Graph].* Online. [Accessed: November 20, 2024]. URL: https://www.statista.com/statistics/1303742/number-of-reported-attacks-against-healthcare-worldwide/ (cit. on p. 1).

Zhang, Min and Juntao Li (2021). "A commentary of GPT-3 in MIT Technology Review 2021". In: *Fundamental Research* 1.6, pp. 831–833. ISSN: 2667-3258. DOI: https://doi.org/10.1016/j.fmre.2021.11.011. URL: https://www.sciencedirect.com/science/article/pii/S2667325821002193 (cit. on p. 24).

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

Eidesstattliche Erklärung
laut §5 der Promotionsordnung vom 15.06.2018

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

Ort, Datum

Simon Benedikt Weber

This DVD contains:

- A *PDF* version of this thesis
- All LaTeXand grafic files that have been used, as well as the corresponding scripts