University of Twente

Computer Science
Services, Cybersecurity and Safety

Master Thesis

Unboxing Security Analytics:

Towards Effective Data Driven Security Operations

Name:
Herman Slatman

Supervisors:
Dr. R. B. N. Aly
Dr. M. H. Everts

Thesis submitted in June 2016

This thesis is dedicated to my parents, without whom I would
not have been able to initiate, persevere through
and successfully complete the education I
received up to this point in my life.

ABSTRACT

Security Operations Centers (SOCs) play a central role in protecting organizations from diverse threats targeting their primary business processes. It is their mission to detect, analyze, respond to, report on and prevent security incidents. Despite substantial investments in preventive and detective security controls, adversaries still manage to remain undetected for prolonged periods of time which can result in a security breach. SOCs face hard times when protecting their constituencies due to diverse causes. This thesis addresses these difficulties by introducing a holistic approach to security operations: *Data Driven Security Operations*.

We first performed an investigation of the challenges SOCs face these days based on gray literature. We categorized the resulting challenges into four main categories: an increasingly complex IT environment, limited business alignment, ever - evolving adversaries and corresponding attacks, and finally, inadequate resources with respect to people and technology. A description of each of these categories and its associated elements are part of the problem analysis and formalization.

We address the challenges by presenting a holistic approach to security operations: the conceptual model for *Data Driven Security Operations*. The model consists of the following six facets: *Situational Awareness*, *Threat Intelligence*, *Detection Methods*, *Response & Investigation*, *SOC Staff* and *SOC Infrastructure*. All six facets revolve around data and together they show how people, processes and technology are all crucial elements to perform security operations driven by data and analysis thereof.

We also created an instantiation of the conceptual model for *Data Driven Security Operations*. SOCs can use it to assess their current status with respect to the six facets. Performing the assessment increases the tangibility of the model, lays the foundation for discussing the effectiveness of the SOC and provides recommendations for improvement.

Both the model and the instantiation were evaluated with five professionals. Although the interviewees indicated that they liked the instantiation, several improvement points were identified. The conceptual model itself was received positively.

"If you think technology can solve your security problems, then you don't understand the problems and you don't understand the technology."

— Bruce Schneier

ACKNOWLEDGEMENTS

This thesis is the result of my graduation project at the University of Twente as part of the Kerckhoffs Institute. Completing this final step to attain my masters degree would not have been possible without the support of many people.

First I would like to thank my supervisors, Dr. Robin Aly and Dr. Maarten Everts for supporting me during the time of performing this thesis. They provided me with useful feedback and advice to get me back on and keep me on track of what I experienced as a rollercoaster ride.

I would also like to thank the interview participants, and the organizations they work for, for devoting some of their time to discuss my work. Their contribution was not only critical to the completion of my work, but also inspiring and fun.

Finally I would like to thank my family, friends and other people for supporting me in many ways whenever necessary, helping me cope with the obstacles I faced and providing me with opportunities to disperse my attention from the subject at hand.

Enschede, 6 June 2016

Herman Slatman

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

This chapter introduces the context of this research and the problem we address in Sections 1.1 and 1.2 respectively. The research questions and a summary of the research strategy are described thereafter in Sections 1.3 and 1.4. The structure of this thesis is provided in Section 1.5.

1.1 CONTEXT

According to the yearly report on data breaches by Verizon, the number of confirmed occurrences of data breaches grew with 55% between 2013 and 2014 [140, 141]. Despite substantial investments in preventive security controls, rated by Hewlett-Packard (HP) at 86% of the total budget available for security [117], organizations are still getting breached on a daily basis [141]. Employing preventive controls and relying solely on those is clearly not sufficient to stop attackers in their tracks.

To make matters worse, current detective security controls fail to detect persistent threats resulting in a median detection time of 205 days [122]. In exceptional cases threats manage to stay below the radar for over 15 years before being detected [135]. The massive number of malware instances reported by Dell [114], to a very large extent unique to an organization[1], further strengthens the need for more advanced detection methods. Current detective security controls, typically implemented in systems like Intrusion Detection Systems (IDSs) and usually based on either predefined rules or anomaly detection [35], also seem to fall short in practice nowadays.

In the continuous arms race between attackers and defenders a clear need for a data-driven approach to security has emerged in the minds of defenders [137]. Rudimentary implementations to realize this approach already exist: so-called Security Information and Event Management (SIEM) solutions. A SIEM solution provides organizations with the technological

1 37 million malware instances, of which 70-90% are unique [141], meaning having a distinct hash or signature.

means to aggregate event data produced by various security devices, network infrastructure, systems and applications [119]. Although SIEM technology has matured over the past couple of years, many deployments failed to realize their full potential [111]. Some of the causes include integration of the SIEM in complex Information Technology (IT) environments and technical and operational skill deficiencies within the Security Operations Center (SOC) [70, p. xxv - xxi]. SIEM solutions also rely on pre-defined use cases to be implemented in order to detect threats and it can be hard to acquire the data necessary to operationalize the use cases.

The Security Operations Center (SOC) consists of the people, processes and technology to protect the organization and mitigate risks to critical business assets. It sits at the core of an organization and is responsible for security operations and Computer Network Defense (CND). The team of security analysts is organized to execute a single mission including detecting, analyzing, responding to, reporting on and preventing cybersecurity incidents [106, p. 9]. In order to fulfill their mission, the SOC delivers a number of services and capabilities to its constituency. These capabilities can be categorized in *real-time analysis, intel and trending, incident analysis and response, artifact analysis, tool support, audit, assessments and outreach* [106, p. 19 - 24]. Most of these are increasingly being supported by technology and data. One example is audit logging: logs of human interactions with information systems are analyzed in order to assess who may have had access to certain data. The SOC's mission is largely supported by technology and may fail when analysts do not have access to the right tools and data at the right time and in the right context. This motivates the need for the right technological solutions to be in place within the SOC.

Security Analytics (SA) was coined *the next big thing in IT security* by Network World [126] in May 2013. It was announced as a bridge between current detective controls, such as SIEM, and the possibilities that new data processing technologies provide [13, 14]. In short, we characterize Security Analytics (SA) as *comprising technologies for economically extracting actionable security intelligence from very large volumes and a wide variety of information security data by enabling high-velocity capture, discovery and analysis intended for improving information security management in a highly dynamic IT landscape subject to ever-evolving*

threats[2]. It could provide SOCs with the means to effectively collect all sorts of security relevant data and process it in an efficient and manageable way. Furthermore, it could also provide analysts with the technologies they need to extract actionable intelligence from the collected data, accomplished by performing advanced pattern mining and ad-hoc analysis over complete data sets. The newly acquired intelligence can lead to improved situational awareness, a shorter time between breach and detection, improved investigation capabilities and reduced residual risk. The ultimate goal of SA could be described as providing organizations with the means to analyze their data to uncover potential security incidents in a more timely manner and to become more nimble in an ever-changing IT landscape that is constantly under attack by emerging threats.

1.2 PROBLEM STATEMENT

Organizations are facing an increased number of ever-evolving threats targeting their primary business processes. The rapid developments in the current threat landscape introduce complex challenges related to scalable and high speed data integration and analysis, threat detection and incident response, severely taxing operators in the Security Operations Center (SOC). Without adequate technology at their disposal and processes in place, security operators are not capable of protecting their constituency.

By bridging between current detective controls, including Intrusion Detection Systems (IDSs) and Security Information and Event Management (SIEM) systems, and the possibilities provided by new technologies for processing complex data, the concept of Security Analytics (SA) certainly sounds promising. It may provide security operators working within the SOC with the tools they need to perform divergent types of analysis in a more efficient and effective manner and to detect threats earlier.

SA is still an ill-defined concept however, and current solutions marketed as SA are presented like a panacea that solves all of the security challenges an organization faces. Organizations keep buying and deploying *boxes* without being able to assess grounded evidence of the actual performance of these solutions

2 Definition based on the definition of *Big Data* by the International Data Corporation (IDC) [116].

while vendors continue to make exorbitant claims. A complete view of what it takes to deploy and operate SA solutions, as characterized in Section 1.1, within a SOC is missing however, which is the problem we address in this thesis.

1.3 RESEARCH QUESTIONS

The goal of this research is to improve the understanding Security Operations Centers (SOCs) have about Security Analytics (SA) and what it takes to deploy and operate this class of solutions. We argue that an integrated approach consisting of people, process and technology is necessary and we present this holistic approach as *Data Driven Security Operations*. In addition to providing organizations with a complete overview of the concept, they need to be able to assess themselves within *Data Driven Security Operations* to relate the concept to their practice and to increase their understanding of what it means to them specifically. To this end, we will answer the following research questions:

1. What challenges do SOCs face nowadays?
2. What should a SOC take into account to address these challenges?
3. How can SOCs position themselves within *Data Driven Security Operations*?

1.4 APPROACH

The design of artifacts, a conceptual model and associated instantiation in this case, forms the core of this research. Because of this, the relatively new concept of Design Science Research (DSR) was applied to perform the research. Figure 1.1 shows the activities generally involved in DSR [77] and which specific research methods were applied at each step in this thesis. First the research problem and its context were formalized. A conceptual model is proposed as a viable solution for addressing the problem. The design of the conceptual model involved an investigative study of scientific literature, survey results and public material supplied by vendors, including mar-

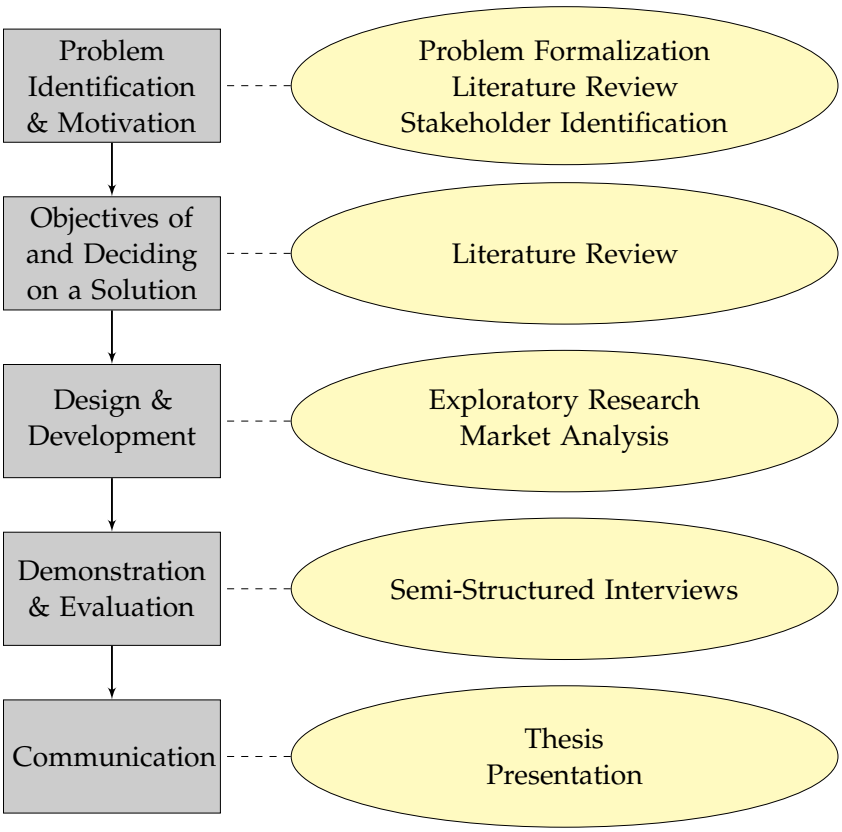


Figure 1.1: The research strategy based on the DSRP. Applicable research methods have been added. Adapted from Peffers et al. [77].

keting material and technical manuals. In addition to the conceptual model, we also created an instantiation of the model which Security Operations Centers (SOCs) can use to position themselves within the conceptual model. A rigorous evaluation of how the research strategy was setup and executed can be seen in Table 7.3, which is based on the guidelines by Hevner et al. [47]. More details about the Design Science Research Process (DSRP) and DSR can be found in Section 2.3 and an extended version of the approach taken for performing this research is described in Chapter 3.

1.5 STRUCTURE

The following chapters address additional background information and how the final research goal was realized. Chapter 2 describes additional background information, including information security, different types of analytics and an expla-

nation of Design Science Research (DSR) and its philosophical assumptions in sections 2.3 and 2.3.1 respectively. It is followed by an extensive description of our research strategy in Chapter 3. Chapter 4 describes our investigation of the challenges that Security Operations Centers (SOCs) face nowadays, which together form the basis for setting the requirements for addressing the problem in Chapter 5. Chapter 6 then describes the process of constructing the conceptual model and an instantiation and shows the final results, addressing the challenges faced by SOCs. The evaluation of the conceptual model, its instantiation and the research strategy are discussed in Chapter 7, which includes a description about how the evaluation was performed. The thesis is wrapped up with conclusions and suggestions for future work in Chapters 8 and 9 respectively.

2 BACKGROUND

In this chapter we present material related to the context of our research. In Section 2.1 we give a description of security operations and Security Operations Centers (SOCs). It is followed by an explanation of analytics in Section 2.2. Finally, we describe the Design Science Research Process (DSRP), a relatively novel approach to and process for designing artifacts using scientific methods, which we used to perform our research.

2.1 SECURITY OPERATIONS

In this chapter we describe what security operations are and elaborate on what Security Operations Centers (SOCs) consist of. We will first give a basic overview of Information Security, because the main reason that SOCs exist in the first place is the need to secure Information Systems (IS) using a centralized approach. The parts thereafter describe what a SOC consists of, how it is organized, the types of work that are typically performed by the SOC staff and how these activities are supported by technology.

2.1.1 An Introduction to Information Security

Many organizations heavily rely on Information and Communication Technology (ICT) in order to execute their primary business processes effectively and efficiently. The types of applications of ICT are ample and diverse, ranging from running administrative software on single desktops, through maintaining mainframes supporting an enormous number of employees and applications, to deploying web applications on a global scale. Disruption of these primary business processes may result in undesired losses to the organization, which is the prime reason to protect these processes and the Information Systems (IS) supporting them.

Information Security is defined as the *preservation of Confidentiality, Integrity and Availability of information [51]*, often de-

scribed as the CIA triad. It is realized by protecting information and information systems from unauthorized access, use, disclosure, disruption, modification and destruction [87]. In addition to the security goals defined in the CIA triad the definition of Information Security can be extended with the properties described by Cherdantseva and Hilton [18]. They describe authenticity & trustworthiness, accountability, auditability, non-repudiation and privacy as additional properties.

The goal of Information Security as a whole is to ensure that security incidents from the past cannot occur (again) within an organization or will at least have a lower impact when they do occur. This goal is realized by establishing, implementing, monitoring and reviewing a suitable set of security controls, the process of which is supported by development of a Information Security Management System (ISMS) according to the specification in ISO 27001 [53]. The process described in ISO 27002 provides a systematic approach to develop and maintain the ISMS [52]. The ISMS in turn provides a holistic view of the risks the organization is subject to by taking into account all aspects of Information Security. We describe *threats*, *threat agents*, *assets*, *vulnerabilities*, *countermeasures* and *risks*, the aspects to consider in Information Security, in the following subsections. The relationships between all of the aspects are shown in Figure 2.1.

Threats and Threat Agents

A comprehensive definition of a threat in an Information Security context is given by the National Institute of Standards and Technology (NIST) [87].

A threat is:

> any circumstance or event with the potential to adversely impact organizational operations (including mission, functions, image, or reputation), organizational assets, individuals, other organizations, [...] through an information system via unauthorized access, destruction, disclosure, modification of information, and/or denial of service.

Threats can either be accidental or intentional. Examples of the former include natural disasters and technical malfunctions, whereas the latter can be exemplified by criminal organizations and malicious insiders. We focus on the latter in this thesis.

Threats can manifest themselves in many forms, illustrated by the European Union Agency for Network and Information

Security (ENISA) *threat landscape* report that was published in January 2015 [127]. The report lists 15 threats that organizations faced during 2014 in order of occurrence together with their trends since December 2013. Malicious code, web(application) attacks and *botnets* were the threats that manifested themselves the most during the analysis period. In addition to the trends in the current Information Technology (IT) landscape, the report also describes the trends that threats are subject to in emerging areas which include *Mobile Computing (MC)* and the *Internet-of-Things (IoT)*.

A *threat agent* (or *threat source*) is an entity that can be categorized by intent and method that can manifest an intentional or accidental threat to happen [87]. Casey, Koeberl, and Vishik [15] present a threat agent classification that can be used to classify threat agents. They discuss the Intel Threat Agent Library (TAL) [110], which addresses the historical problems of information about threats being fragmented, sensationalized and lacking standard definitions. The TAL defines a total of 21 different threat agents, classified as either being non-hostile or hostile and ranging from employees to government spies. Identification of threats is based on the following 8 attributes: *intent*, *access*, *outcome*, *limits*, *resources*, *skills*, *objective* and *visibility*. Having a clear view of which threat agents are targeting the organization is an important piece of information to take the right decisions when allocating security budget.

Threat analysis.is a method that can help to determine which threats an organization faces or will face in the future. It can be described as *the examination of threat sources against system vulnerabilities to determine the threats for a particular system in a particular operational environment [90]*. It is an important part in the process of threat modeling that results in a model that can be used to evaluate the security posture of a system [92] or that can help when performing security tests against applications [94] or networks [69].

Vulnerabilities and Countermeasures

Perfect security is impossible to achieve and as a consequence vulnerabilities will always exist in information systems, security procedures, internal controls or software & hardware implementations [87]. These vulnerabilities can be triggered or exploited by a threat which results in a security incident. An important step in selecting the right countermeasures is determining what vulnerabilities an organization is susceptible

to. Countermeasures are the actions, devices, procedures, techniques or other measures that meet or oppose (i.e., counter) a threat, a vulnerability, or an attack by eliminating or preventing it, by minimizing the harm it can cause, or by discovering and reporting it so that corrective action can be taken [74].

Assets and Risk

Risk is the net negative impact of the exercise of a vulnerability, considering both the probability and the impact of occurrence [89]. The proper management of risks within an organization is a crucial step in the process of getting better at managing IT related risks. Security Risk Management (SRM) is a process that takes into account the identification of risks, assessments of risk and taking actions to reduce the risks an organization is subject to [89]. The ultimate goal of SRM is to improve the security of business critical IT systems within the organization, accurately estimating and budgeting IT security related purchases and supporting management in decisions regarding authorization of IT systems.

A so-called Security Risk Assessment (SRA) can be conducted to determine the risks an organization is subject to. A SRA is an objective analysis of the effectiveness of the current security controls that protect an organization's assets and a determination of the probability of losses to those assets [63]. The goal of an SRA ultimately is to reduce *operational cyber security risks*, which have been defined by Cebula and Young [16] as:

> The operational risks to information and technology assets that have consequences affecting the confidentiality, availability or integrity of information or information systems.

This goal can be realized by first getting a complete view of the assets including their value to the organization and an assessment of the current threat environment. Based on the value of an asset, the threat environment and security controls currently in place, the likelihood and impact of the asset being compromised can be calculated. Recommendations regarding prioritization of and what security controls to implement can be deducted from the results of the analysis. The report by Cebula and Young [16] also presents a taxonomy of the operational cyber security risks that can be used to help identify the operational cyber security risks an organization may be subject to. Cyber risks can be categorized into four main classes

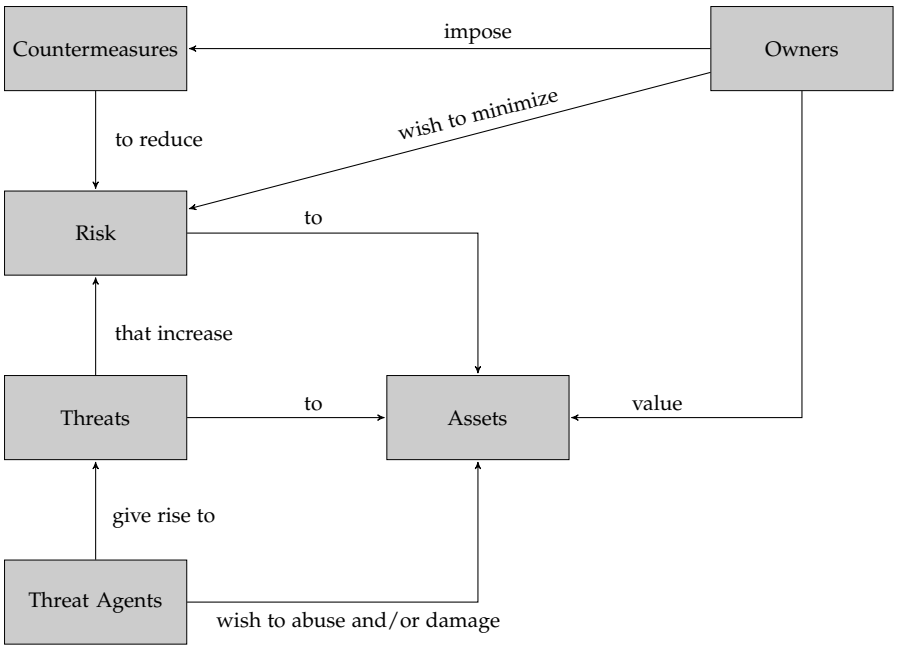


Figure 2.1: Relationships among the elements of risk. Adapted from ISO 15408-1:2009 [50].

as described in the report: *actions of people*, *system and technology failures*, *failed internal processes* and *external events*. The taxonomy also relates to other standards that describe cyber risk, including the Federal Information Security Management Act (FISMA) [1], NIST 800-53 [87] and the Operationally Critical Threat, Asset, and Vulnerability Evaluation (OCTAVE) methodology [11]. Continuous awareness of and adaptations to the SRM framework are necessary in order to account for changes in the organizations *IT* architecture and threat landscape.

2.1.2 The Security Operations Center

Organizations are increasingly being targeted by various threats. To get a better understanding of and counter these threats they can deploy a SOC, which sits at the core of an organization in terms of information security and provides Computer Network Defense (CND). The SOC consists of the people, processes and technology to protect the organization and mitigate risks to critical business assets. Its mission can be described as *detecting, analyzing, responding to, reporting on and preventing cybersecurity incidents* [106, p. 9]. As part of this mission the SOC delivers a number of services and capabilities which can be categorized as *real-time analysis*, *intelligence and trending analysis*, *incident anal-*

ysis and response, *artifact analysis*, *tool support*, *audit*, *assessments* and *outreach* [106, p. 19 - 24]. Most of these services, capabilities and, consequently, the SOC's mission, are increasingly being supported by technology and data. Having access to the right tools and data at the right time and in the right context is thus critical for the SOC staff to successfully fulfill their mission.

In the following subsections we describe in more detail what the SOC and its members do and what type of technology is typically deployed to perform these activities.

Security Operations, Services and Capabilities

Despite the fact that a clear definition for the SOC exists (see: Section 2.1.2), this does not mandate what a SOC should consist of and what it should do in reality. The number and types of services and capabilities offered by a SOC can differ considerably across constituencies, which are ultimately what make up the SOC. These services and capabilities can be logically categorized in three categories: reactive, proactive and quality management services [103].

Reactive services are initiated after a trigger such as an alert or a request. Examples of these include the report of a compromised host by an employee, a publication about a severe vulnerability or an alert generated by a Security Information and Event Management (SIEM) system. Depending on the criticality of the alert or request, the SOC has to triage the alert or escalate it to a security incident. Security incidents have to be analyzed, contained, eradicated and recovered from [22]. This also includes acquiring forensic evidence and analysis thereof, such as performing reverse engineering of malware [103] and disk or memory forensics [106]. Active coordination of all the aforementioned services is also part of the mission of the SOC.

Proactive services are aimed at preparing and securing the IT environment before an attack occurs. Security monitoring and intrusion detection are important tasks in this category [103]. Performing security audits can also be part of the tasks to be executed by the SOC. These include infrastructure reviews, active and passive scanning, vulnerability assessments and penetration tests [106]. Reporting on the current status of information security, the creation of policies and workflows and playbooks in preparation of security incidents also belong to the proactive services [22]. These days SOCs are also increasingly responsible for tracking what types of threats exist, how these evolve over time, what Tactics, Techniques and Procedures (TTP) are rele-

vant when assessing these specific threats and which are most important to focus on [106].

Services in the quality management category can be quite varied and are not necessarily performed only by the SOC or its sole responsibility. Often the members of the SOC have to cooperate with other business departments, such as IT operations and the legal department. Development and configuration of software and hardware used for performing the reactive and proactive services is only one of their tasks [106], but critical nonetheless. Because of their broad and deep subject matter expertise the SOC can also provide input to risk analyses and business continuity planning, perform product evaluations and security consulting services and deliver security awareness training and education [103].

The number and variety of tasks that a SOC may perform, some of which we have mentioned above, is indeed daunting. A SOC should not focus on performing all of these all at once, but gradually increase the number of capabilities and services offered and only when enough resources are available. It is better to get good at only a number of the tasks than to be mediocre or bad at all or some of them [103, 106]. Another factor that plays a role in what specific services and capabilities are offered by a SOC is the organization of the SOC, which we describe next.

SOC Organization

As described before, the organization of a SOC largely depends on what services and capabilities are offered. Usually the SOC consists of several tiers in which specific duties are performed [139]. The first tier may handle most of the events and alerts that can be handled fairly quickly during the initial triage. When a certain alert needs more time to be addressed, second tier analysts can take over the analysis, which can optionally be followed by tier three analysis. In some situations more specialist knowledge is necessary, such as for performing memory, disk or malware forensics. Other services and capabilities are less reactive, and may need constant or regular attention, which can be provided by specific SOC members or during times that require less ad-hoc analyses.

Below we list several possible organizational models a SOC can operate in, but as with any organization, many more forms exist in the real world. The examples we provide vary in the number of services and capabilities offered, level of authority,

number of employees and the number of connections to external entities.

- A group of IT operators, potentially dispersed within an organization, operating under a single person to address security alerts and incidents.
- A small to medium, internal SOC, offering security monitoring and triage in two distinct tiers, but outsourcing the handling of security incidents to external parties having the right capabilities.
- A coordinating SOC managing multiple locations of the same constituency, which can be geographically dispersed, for example. Its role is to provide strategical and operational guidance and is less focused on the tactical level.
- A full-fledged internal SOC offering most of the services and capabilities itself. In some cases the expertise from external parties can be called upon. This type of SOC is typically operated in larger organizations.
- When offering all or some of the aforementioned services and capabilities as a service, the SOC can be characterized as a Managed Security Services Provider (MSSP). Many smaller organizations are starting to realize the value of security monitoring and incident response, but do not have the resources to setup their own SOC and will procure certain services via an MSSP.

These are only some of the possibilities of how a SOC can be organized with respect to its constituency. The SOC itself can also be organized in different ways, such as a flat and wide layout, distributed over several parts of the constituency or hybrid forms of these [118].

SOC Infrastructure

Without any infrastructure to support all of the services and capabilities offered, the SOC would not be able to operate effectively. What the SOC needs in terms of technology and how these systems should interact heavily depends on the services and capabilities being or going to be offered, but also on the scale of its constituency and legal requirements. It may also be the case that the SOC heavily depends on approval or cooperation from other departments within the constituency, such

as IT operations, when setting up the infrastructure to deliver services and capabilities. An example of this is setting up and configuring logging on a proxy server, for which approval may be necessary.

There exist numerous technologies that enable the SOC to perform its job. The first element is the collector platform, which consists of several types of sensors and storage nodes. Sensors are entities that read aspects from one or more IT systems, extract knowledge and generate data. Three characteristics that define what data can be generated by a sensor are its *vantage*, *domain* and *action* [26]. The vantage point defines what part of an IT environment a sensor is able to sense, such as a single networking link, a single host, or all of the network traffic. What kind of information a sensor can provide is determined by its domain, such as networks, hosts or services. The third characteristic that defines a sensor is what action it performs upon sensing its environment, which can consist of simply reporting the data, signaling an event based on aggregated data, actively acting upon its environment or a combination of these. Some examples of sensors in the information security domain include anti-virus (AV), (application) firewalls, Intrusion Detection Systems (IDSs), Intrusion Prevention Systems (IPSs) and network taps.

All of the data that is produced by the sensors has to be collected and stored. There are many aspects that one has to take into account when creating a (centralized) storage for security data. The first is the format of the data that is going to be collected. Many types of sensors provide the notion of a log: a collection of event records, that together describe the sequential occurrence of events within an environment [8], such as state changes and user interaction. Good logs contain complete information that describe what, where, when, and why a certain event happened and who is involved [21, 26].

Another type of data to be collected includes alert data, which may seem similar to events, but should be seen as a separate type of data that demands actions to be taken. SOCs can also benefit a lot from looking at network capture data with various fidelity, ranging from packet meta data, flow data to full packet capture, to improve their visibility into complex networks [79].

The SOC infrastructure should take care of secure and reliable transport and storage of various formats of data originating from a diverse set of sensors. During the past couple of

years the volume of data collected from the environment for security monitoring and incident response has grown dramatically, which resulted in the need for scalable technologies to be in place for data storage and processing. The prime example of secure, centralized (log) storage in recent years is the SIEM, which provides analysts with a view of the IT environment as reported by all the logs that are configured to forward their contents to the SIEM system. In addition to providing a centralized log storage facility, SIEM systems also can be used to create aggregated alerts resulting from running correlation rules on logs, alerts and network data. Depending on the quality of the correlation rules, these alerts ideally are a strong indicator of a security incident that has taken place which needs triage. In addition to providing security analysts with the necessary data for performing real-time security monitoring, the (centralized) storage of logs, alerts and networking data is also helpful when performing incident response and forensic investigations.

2.2 ANALYTICS

Analytics is the extensive use of data, statistical and quantitative analysis, explanatory as well as predictive models and fact-based management to drive decisions and actions [32]. Within organizations it facilitates the realization of business objectives through reporting of data to analyze trends, creating predictive models to foresee future problems and opportunities and analyzing/optimizing business processes to enhance organizational performance [33]. The latter is measured by Key Performance Indicators (KPIs).

Business Intelligence (BI) is a set of technologies and processes that use data to understand and analyze business performance [32]. Business Analytics (BA) comprises both analytics and BI. The term has been used in organizations to describe the usage of information technology to gain insight from data. It applies to software products, analytics solutions areas, consultancy services, outsourced business processes and hardware [65].

There are several ways to perform data analysis, each of which has its own advantages, disadvantages and consequences regarding the type and quality of insights that can be delivered. What kind of analysis to perform depends on what an organization actually wants to know and what it needs. The needs

of organizations can be categorized along five axes: *access to information*, *insight*, *foresight*, *business agility* and *strategic alignment* [65].

The following subsections describe the three types of analytics that are recognized. Each successive type described in the following subsections is increasingly sophisticated in terms of analysis approach and typically results in greater insights for improved decision making.

2.2.1 Descriptive Analytics

Descriptive Analytics (DsA) is the start of the process of gaining new insights from data. It is a set of technologies and processes used to understand and analyze the current business performance [65]. The main objective is to find out what has happened in the first place, why that event could take place and what is happening right now in order to learn from these events and identify business opportunities and problems [33]. Typical application of DsA and BI software results in dashboards and sales reports, often including visualizations of the available data. DsA is useful to gain insight from past events and provides organizations with a single view of the past, which allows them to focus on the present and future [65].

2.2.2 Predictive Analytics

Predictive Analytics (PdA) is a step up from DsA and is about turning data into actionable and valuable intelligence. It is about taking historical data and understanding thereof to predict future events [65]. It can be applied to both offline (e.g. determining similar groups of customers for targeted mail) and real-time processes, such as detecting fraudulent transactions. Various techniques can be applied using PdA, including statistical analysis, predictive modeling and simulation as well as forecasting.

2.2.3 Prescriptive Analytics

Once an organization knows what happened in the past and can predict the future to a certain extent, it can start to think about determining what the right course of action is. Prescriptive Analytics (PsA) plays a major role here as it uses data and

algorithms to determine a set of alternative decision options based on business objectives, requirements and an arbitrary number of constraints. The decision options include the possible implications of the option being taken which can be evaluated by an analyst in order to take the best option for the organization at that point in time. Techniques applied in PsA include *optimization*, *expert systems* and *simulation*.

2.3 DESIGN SCIENCE RESEARCH AND PROCESS

Research in the Information Systems (IS) discipline can be described as a form of research where theory from other disciplines is applied in a practical context [77]. Amongst others, these disciplines include Computer Science, Social Sciences and Mathematics. Theories from these disciplines are applied to solve problems that organizations are struggling with at a technological level, which revolve around Information Technology (IT) systems a lot of the time.

The term Design Science (DS) was first coined by Buckminster Fuller [9] who defined it as a systematic form of designing. This concept was further elaborated on by Simon [85], who persuaded the creation and development of methodical and formalized methodologies for design. Research in IS design has seen much debating on how research should be conducted from a philosophical view, relying on epistemology, theoretical perspectives and methods from natural sciences research [29, 102], which are aimed at trying to find the truth. In contrast, Design Science Research (DSR) tries to determine what is effective instead [47].

Hevner et al. [47] introduce seven guidelines for assessment of research conducted using DSR, which are shown in table 2.1. They also describe the construction of artifacts, including constructs, models, methods and instantiations. Artifacts rarely are full instantiations of information systems, but can be seen as the innovations that define ideas, practices, products and new businesses [36].

Table 2.1: DSR guidelines by Hevner et al. [47].

Guideline	Description
Design as an Artifact	DSR must produce a viable artifact in the form of a construct, a model, a method, or an instantiation.
Problem Relevance	The objective of DSR is to develop technology-based solutions to important and relevant business problems.
Design Evaluation	The utility, quality, and efficacy of a design artifact must be rigorously demonstrated via well-executed evaluation methods.
Research Contributions	Effective DSR must provide clear and verifiable contributions in the areas of the design artifact, design foundations, and/or design methodologies.
Research Rigor	DSR relies upon the application of rigorous methods in both the construction and evaluation of the design artifact.
Design as a Search Process	The search for an effective artifact requires utilizing available means to reach desired ends while satisfying laws in the problem environment
Communication of Research	DSR must be presented effectively both to technology-oriented as well as management-oriented audiences.

2.3.1 Philosophical Assumptions of Design Science Research

Philosophical assumptions determine to a large extent how novel research is performed. They consist of ontology, epistemology and axiology. Vaishnavi and Kuechler [98] define these three concepts as listed below:

- **Ontology** is the study that describes the nature of reality. What is real and what is not, i.e., what do we think is real? What is a fundamental element, and what is of derivative nature?
- **Epistemology** is the study that explores the nature of knowledge. Where does our knowledge come from and how certain can we be about the knowledge that we have?
- **Axiology** is the study of values. What values does an individual or group hold, why and what creates value for an individual or group?

Vaishnavi and Kuechler [97] also show that DS is based on multiple, contextually situated, alternative world-states. Knowledge is acquired by way of creation within a certain context. Circumscription of knowledge is employed to increase understanding, meaning and certainty. The creation of one or more artifacts, continuous improvement thereof and understanding the problem, contribute value. Development is the core of the methodology employed in DSR. Table 2.2 shows the relation between philosophical assumptions of DSR and includes two other research perspectives for reference.

Peffers et al. [77] were the first to develop a mental model and conceptual process for carrying out DS in the IS discipline. They present the Design Science Research Process (DSRP), which consists of six activities and is based on existing scientific literature on DS in the IS discipline. Their approach is consistent with the DS processes in other disciplines, provides a nominal process for conducting research and presents a mental model for what the output of research conducted using DS looks like. A salient detail: the resulting DSRP was developed using the DSRP itself.

2.3.2 Problem Identification & Motivation

The first activity is to define a specific research problem and explain the value of a solution for the problem. A justified valua-

Table 2.2: Philosophical assumptions of three research perspectives. Adapted from Gregg, Kulkarni, and Vinzé [44] by Vaishnavi and Kuechler [97].

	Research Perspective		
Basic Belief	Positivist	Interpretive	Design
Ontology	A single reality. Knowable, probabilistic	Multiple realities, socially constructed	Multiple, contextually situated alternative world-states. Socio-technologically enabled
Epistemology	Objective; dispassionate. Detached observer of truth	Subjective, i.e., values and knowledge emerge from the researcher-participant interaction	Knowing through making: objectively constrained constructed within context. Iterative circumscription reveals meaning
Methodology	Observation; quantitative, statistical	Participation; qualitative. Hermeneutical, dialectical	Developmental. Measure artifactual impacts on the composite system
Axiology	Truth: universal and beautiful; prediction	Understanding: situated and description	Control; creation; problem-solving; progress (i.e., improvement); understanding

tion of a solution helps the researcher and target audience to actively pursue the solution. A complete understanding, gained by e.g., systematic inquiry and deep analysis, of the problem and its complexity is necessary to successfully complete this phase.

2.3.3 Objectives of a Solution

A solution always has the objective of solving a specific problem. This activity includes specifying what a possible solution should accomplish, i.e., its requirements. The extent to which the solution solves a problem can either be quantitatively or qualitatively (or a combination of both) be determined along several axes. Knowledge about the problem and existing solutions is a necessity.

2.3.4 Design & Development

This activity revolves around determining the desired functionality of an artifact. After the desired functionality has been determined, the architecture of the solution and actually creating the artifact is realized by moving from the objectives of a solution towards design of solution elements. This step requires knowledge about possible solution elements and how to apply them.

2.3.5 Demonstration

The produced artifact is demonstrated in order to determine its efficacy. There are several ways to determine the efficacy, including experimentation, simulation or a case study. It requires knowledge about the artifact and a sufficiently sized group of stakeholders in order to complete successfully.

2.3.6 Evaluation

As is the case for any scientific research, the evaluation of the results of the conducted work is a crucial component of the research process [47]. For the constructed artifact this may include observing or measuring how well it works as a solution to the problem that was identified in the first activity. It requires quantitative or qualitative analysis of measures relating

to the objectives of the solution. Results of the evaluation can be used to improve the created artifact in an iterative approach. Table 2.3 lists various evaluation methods that can be applied in DSR.

Evaluation Methods in Design Science Research

As described in section 2.3.6, evaluation is a critical component of any research process. There has been much debate on what evaluation methods can be used effectively in DSR. Table 2.3 shows there are numerous evaluation methods that can be applied. Some frameworks for selecting and validating which evaluation method to choose have been proposed in the past [23, 78]. Venable, Pries-Heje, and Baskerville [99] extended the framework by [78] and used it to construct a *Four-Step Method for DSR Evaluation Research Design*.

Venable, Pries-Heje, and Baskerville [99] adapted the 2x2-matrix presented by [78] to construct a DSR Evaluation Strategy Selection Framework. This resulted in two matrices, the first of which describes a mapping from relevant contextual aspects to the framework by Pries-Heje, Baskerville, and Venable [78]. These aspects include the purpose of the evaluation, the type and characteristics of the evaluand[3] and specific goals of the evaluation itself. The second matrix ([99, fig. 3]) describes a mapping from the framework by Pries-Heje, Baskerville, and Venable [78] to existing evaluation methods.

Focus Groups for Evaluation of Design Science Research

In the social sciences, Focus Groups (FGs) are a widely used research method for evaluation [42]. It can be described as an interview with multiple respondents participating and interacting with each other, discussing a specific topic [57]. They are designed to collect data through group interaction [71]. As a research method, they are located between participant observation and semi-structured interviews [42]. They are aimed at getting an understanding of the topic from the different perspectives participants have by analyzing the data that results from the discussion. Gibson and Arnott [42] argue that because of the interaction between participants, they can become more creative and can address the topic in greater depth when compared to normal interviews: novel ideas may emerge from the interaction. Two main types of FG exist: Exploratory Focus

3 The object or artifact under evaluation

Table 2.3: DS evaluation methods by Hevner et al. [47].

Method	Examples
Observational	*Case Study*: Study artifact in depth in business environment
	Field Study: Monitor use of artifact in multiple projects
Analytical	*Static analysis*: Examining structure of artifact for static qualities (e.g. complexity)
	Architecture analysis: Study fit of artifact into technical IS architecture
	Optimization: Demonstrate inherent optimal properties of artifact or provide optimality bounds on artifact behavior
	Dynamic analysis: Study artifact in use for dynamic qualities (e.g. performance)
Experimental	*Controlled experiment*: Study artifact in controlled environment for qualities (e.g. usability)
	Simulation: Execute artifact with artificial data
Testing	*Functional (Black Box) Testing*: Execute artifact interfaces to discover failures and identify defects
	Structural (White Box) Testing: Perform coverage testing of some metric (e.g. execution paths) in the artifact implementation
Descriptive	*Informed Argument*: Use information from the knowledge base (e.g. relevant research) to build convincing argument for the artifacts utility
	Scenarios: Construct detailed scenarios around the artifact to demonstrate its utility

Groups (EFGs) and Confirmatory Focus Groups (CFGs) [95], the former of which is used during the design cycle whereas the latter is used at the final evaluation stage of the research.

Tremblay, Hevner, and Berndt [95] present a couple of reasons why focus groups are an adequate method for evaluation in DSR based on [88, p. 42]:

1. **Flexibility** is provided by the open format of the FG.
2. **Rich data** can be gathered from the natural interaction among participants in the FG.
3. **New ideas** and opinions emerge from the interaction among participants which usually stay uncovered in normal interviews.
4. **Direct interaction** between researcher and participants allows him to clarify elements of the artifact that remain unclear from the initial explanation.

Figure 2.2 shows the steps typically taken in succession when performing a FG. It is adapted from Tremblay, Hevner, and Berndt [95, fig. 1], which is based on [71, 88]. Performing a FG is certainly not an easy task to accomplish, and thus Gibson and Arnott [42] have created six general guidelines for performing these. Their guidelines are listed in table 2.4.

Strengths and Weaknesses of Focus Groups

Earlier we listed several properties of FGs that render it an adequate evaluation method for DSR. These included a flexible approach, rich data gathering, naturally emerging ideas and direct interaction between participants and researcher(s). The fact that several participants give their opinion on the utility, quality, and efficacy of a design artifact during a single session makes FGs a time-effective method for performing evaluations. A natural tendency to consensus amongst the participants relieves the researcher from extracting this consensus himself.

The selection of a facilitator and facility, correct recruitment of participants and a well-prepared facilitator guide are three key success factors in performing a FG [43]. In general, performing a FG requires rigorous planning time [72]. The fact that the researcher has an active, participatory role during the discussion may introduce a bias towards his opinions [71], but this is true for other qualitative research methods also. Taken

Table 2.4: FG guidelines by Gibson and Arnott [42].

Guideline	Description
Maintain Focus	Focus groups are not random discussions, they can solicit concentrated amount of focused data. Stay on track, plan questions carefully.
Be Selective with Participants and Group Size	Participants are rarely selected randomly. Avoid power differentials between participants. There must be a suitable level of diversity to encourage discussion, however too much will cause conflict amongst participants. Group size will be dictated by the research focus, participant availability, and level of participant involvement in the topic. Six to eight is a good starting point, but accommodate no-shows.
Be Selective with Facilitator	Choose a facilitator familiar with the research area, particularly if it's specialized. They should be personable, and be able to think on their feet. They should guide the group, not control it.
Be Prepared	Carefully plan the facilitator guide, early effort will improve data collection through focused questions. Send any documents early, have spare copies ready on the day. Use fail-safes when technology is involved. Using assistants reduces the researcher's workload, allowing them to focus on key matters. Be familiar with the literature on focus groups, learn from the mistakes of others.
Allow Flexibility	Adapt to change, allow participants to take discussion in useful directions, the facilitator guide should allow for this. Pursue unanticipated questions or comments. Remove questions already covered between groups.
Take a Pragmatic Approach to Analysis	Choose a suitable analysis method. Ensure the analysis approach enables effective data capture, and data is not under-analyzed. Encourage observers to take notes during the sessions. Non-verbal data can be useful, such as laughter, direction of conversation, and facial gestures; video recording sessions aids this.

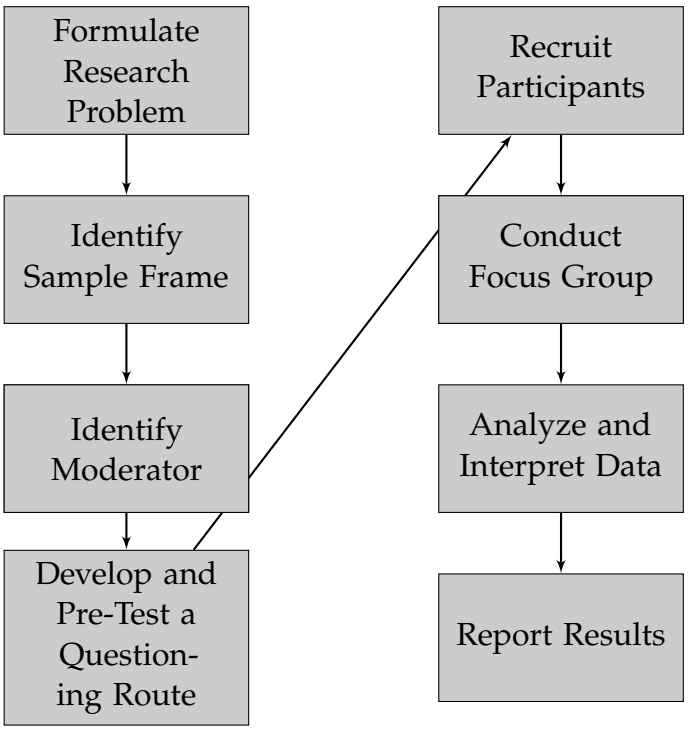


Figure 2.2: Steps to be performed in an evaluation by FG. Adapted from [95, fig. 1].

together, these factors increase the risk of insufficient or inadequate results emerging from the evaluation.

Interviews for Evaluation of Design Science Research

The interview is a qualitative research method that has been used extensively in the social sciences [71]. It can also be used in DSR for data collection at several stages of the research process [57], including during the evaluation stage. Interviews can be structured, semi-structured and unstructured, listed in order of ascending amount of control the researcher can exercise on the interviewee. Semi- and unstructured interviews are to be preferred when investigating complex issues, because of the fact that interviewees can more easily express their ideas in those settings [57]. A one-on-one interview can be conducted via several media, including face-to-face, telephone and Internet-enabled communication [30, p. 179]. Before performing the interview a planning is needed to create a set of adequate questions, invite participants and introduce them to the research.

Strength and Weakness of Interviews

Performing an interview generally results in richer data when compared to conducting ordinary questionnaires, because more complex issues can be discussed and elaborated on [57] and the researcher has more control over what can be discussed [71]. An interview also allows for discussing novel ideas and artifacts [30, p. 179]. The time planning and preparation for an interview takes is likely to be much smaller than for preparing FGs. On the other hand, data analysis may take more time [57] because of the number of interviews necessary to perform. The data collected may not represent the truth because of a possibly limited, inaccurate or incomplete view an interviewee has of the subject [30, p. 179]. The quality of the data collected by interviews can also vary because of differing perspectives and knowledge level of the interviewees [42]. Additionally, the researcher has to take care not to introduce (personal) biases during the interview and the analysis thereof which will likely affect the end results [30, p. 179].

2.3.7 Communication

The final activity in DSR is publishing the final result(s) of the conducted research. It basically describes how the previously listed activities were performed. It also includes an assessment according to the guidelines listed in table 2.1 by Hevner et al. [47]. The results have to be communicated in a clear and effective manner for different types of professionals, including engineers and management, for example.

3 RESEARCH STRATEGY

We employed Design Science Research (DSR) as a guide when constructing the strategy for this research. This chapter describes how DSR was applied in our research, starting with which research methods were applied during each of the core activities of DSR. The six core activities, Problem Identification & Motivation, Objectives of a Solution, Design & Development, Evaluation (usually combined with Demonstration) and Communication [77] and the accompanying research methods, are shown in Figure 3.1. The following subsections describe the research methods that were applied at each step and why these were chosen.

3.1 PROBLEM IDENTIFICATION & MOTIVATION

In Section 1.2 we already stated that Security Operations Centers (SOCs) face many obstacles when defending their constituencies from threats. We performed an extensive literature survey to get a more complete view of the difficulties and their causes to get a better understanding of them. Many security solution vendors are offering many different types of solutions claiming to solve some or all of the SOC's challenges, but their products are hard to assess with respect to their effectiveness. We observed that many of the difficulties faced by SOCs involve data and analysis thereof in some way, but simply deploying solutions offered by vendors is not enough. We determined that a complete view of the factors and a holistic approach to apply Security Analytics (SA) within a SOC was missing.

3.2 PROPOSING A SOLUTION

Because we wanted to transfer knowledge and make people understand the bigger picture that we envision, we proposed to

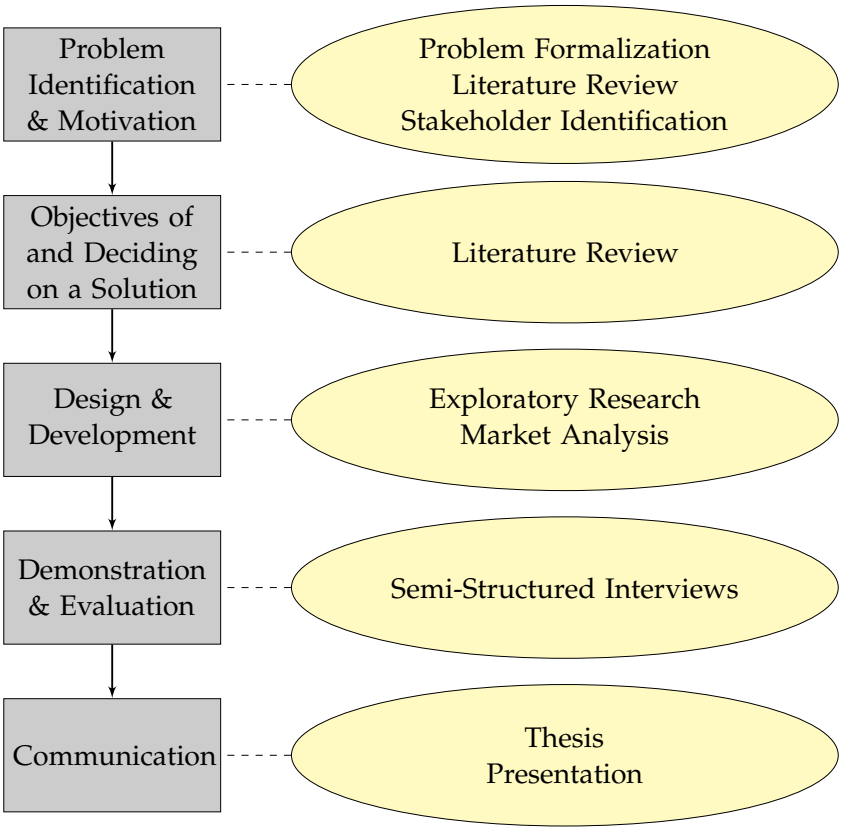


Figure 3.1: Research strategy based on the DSRP, including the research methods that we employ in this research. Adapted from Peffers et al. [77].

create a conceptual model for what we call: *Data Driven Security Operations*. The model had to capture what *Data Driven Security Operations* consists of, which is what our interpretation of implementing Security Analytics (SA) entails, and should give Security Operations Centers (SOCs) and the more general public a complete view of the concept too. In order to provide SOCs a more concrete view of the concept, we also proposed to create an instantiation. Through interaction with the instantiation SOCs will be able to better relate the model to their own practice. The second use of the instantiation is for the SOC to position itself within the field of *Data Driven Security Operations*, i.e. *how well is the SOC doing? How can it become more effective?* We created a set of requirements for both the model and the instantiation that describe what the resulting solution should provide based on the literature review.

3.3 DEVELOPMENT

The first parts of our research required us to do extensive literature research of many subjects, including security operations, analytics and identifying stakeholders. In our research we also assessed the current state of the market to get an idea of the direction security vendors are heading. The sources for these included scientific literature, market research reports, marketing material provided by vendors and general web sources, including weblogs and webinars, conference material, survey results and finally, a limited amount of practical experience with security solutions and analytics platforms. The instantiation is based on the research performed for creating the conceptual model. Its format is based on existing methods and has been constructed according to generally-known formats for questionnaires. Both the conceptual model and the instantiation are constructed while adhering to the requirements that were set in advance.

3.4 DEMONSTRATION & EVALUATION

The demonstration and evaluation of the model and instantiation were combined in this research, because they are closely linked, was a practical thing to do and also results in a more for-

mal evaluation of the artifact [47]. In this thesis we first describe how an appropriate and valid evaluation method was selected, followed by a description of the chosen evaluation method. We also evaluated the entire research strategy using the guidelines published by Hevner et al. [47], which is described in Table 7.3.

3.5 SELECTING AN EVALUATION METHOD

The evaluation of this research involves a socio-technical model as well as an instantiation as artifacts. The type of evaluation can be described as an *implementation evaluation* [104, ch. 5], because it is performed after the artifacts have been constructed and not during the design process. A certain level of conflict may exist, because of the fact that different types of stakeholders are involved in the evaluation and communication of the artifact. Because of time constraints, the amount of data that could be gathered and analyzed is limited. Furthermore, the quality of the gathered data is an important aspect. Because of these reasons, a qualitative evaluation method is preferred.

Given the properties described above the research would fall in the *Ex Post / Naturalistic* quarter of the 2 x 2 - matrix shown in [99, fig. 2]. An *Ex Post* evaluation is performed after an artifact has been created and can be described as an implementation evaluation Because of the socio-technical nature of the artifact and real users being involved, both *Focus Groups (FGs)* and *one-on-one interviews* have been considered as adequate evaluation methods. Although performing a FG seemed more time-effective than performing one-on-one interviews, the risks associated with performing an FG were considered too big to tackle with high confidence. Because of this, semi-structured interviews with experts were selected as the evaluation method for the constructed artifacts. During the interviews the conceptual model was shown to the interviewee first, followed by an explanation of each of the facets. The interviewee then interacted with the instantiation, performing the test, which is followed by a semi-structured discussion. Qualitative data was gathered during all stages of the interviews.

3.6 COMMUNICATION

Our research has been communicated at several stages and abstraction levels. This thesis is the main medium for transferring the knowledge embedded in the conceptual model. We also presented the conceptual model to the participants during each of the evaluations. Finally, participants receive a personalized report produced by the instantiation that contains the description of the conceptual model as well.

4 PROBLEM FORMALIZATION

In this chapter we describe how we formalized of the problem, the first activity in Design Science Research (DSR). This step serves several purposes, including getting a deep understanding of what the actual problem is and what it is caused by. We the justify the need for this research by defining a specific research problem and the value of a solution. We also show stakeholders why this research is necessary.

4.1 PROBLEM ANALYSIS

Security Operations Centers (SOCs) face numerous obstacles when protecting their constituencies and fulfilling their day-to-day operations. We performed a study of the factors that play a role in the impediments they face, which we summarized in the diagram shown in Figure 4.1. Most of the sources we consulted were either whitepapers, gray literature, analyst reports or the results from surveys, because these are more timely and closer to the practice of SOCs than scientific literature on these subjects. We categorized the diverse set of causes of difficulties into four main categories. The following subsections describe the four categories listed below:

(A) An increasingly complex IT environment,

(B) Limited business alignment,

(C) Inadequate resources with respect to people and technology, and finally,

(D) Ever-evolving adversaries and corresponding attacks.

Complex IT Environment

The first main type of difficulty for the SOC is related to the scale and complexity of the IT environment. During the past couple of years the IT environment within organizations has

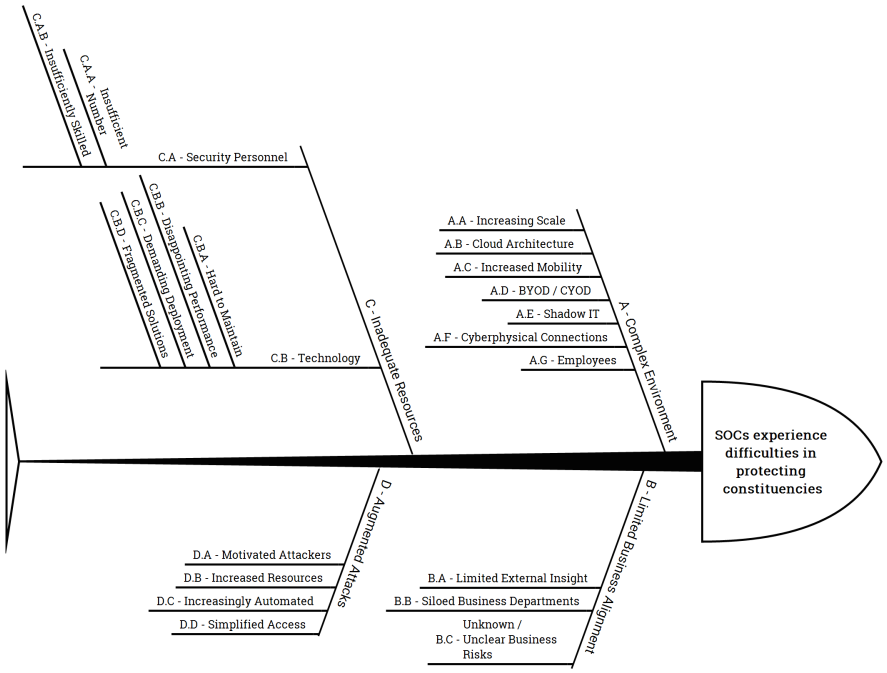


Figure 4.1: Ishikawa diagram showing the causes of difficulties the SOC faces while trying to protect their constituency.

changed tremendously in diverse areas. Due to increased mobility, larger scale IT usage, usage of cloud infrastructure and employees having the ability to bring their own devices to work, SOCs have less visibility into and control over what happens on the network and who has access to which business assets at what time [120, 124, 129, 136, 137, 138]. Employees are also increasingly making use of software not being provided by their employer for various reasons, such as for improving work performance, faster collaboration and better communication. This phenomenon is called *Shadow IT*. The main risks of *Shadow IT* have been identified by Silic and Back [84] and include the decrease or loss of data integrity and information leakage.

The human aspect plays a role in many of the aforementioned factors and is a factor in itself too. The first reason is that making mistakes due to the lack of knowledge, several different kinds of biases and other psychological influences is part of our nature [76]. Secondly, insider threats have grown to become one of the biggest threats to consider within businesses [19, 27, 45]. The insider threat is not only about rogue employees, but also denotes the fact that adversaries are increasingly able to become insiders by infecting employees or obtaining legitimate credentials.

Another development that has attracted interest and will likely cause the SOC hard times is the increase of industrial and ambient networks getting connected to the business network. These so-called cyber-physical systems, which control physical systems using computers, range from the Internet-of-Things to Industrial Control System (ICS), and can include anything from controlling devices to Supervisory Control and Data Acquisition (SCADA). These systems use all kinds of lesser known, sometimes proprietary, protocols and hardware, requiring specific technologies and specialist knowledge to protect [12, 46, 82], which the SOC often does not have available.

Adversaries are also increasingly making use of the intricacies of IT environments and legitimate applications to gain access, move laterally and persist within an organization. Keeping an IT environment up to date and configured appropriately is absolutely necessary to prevent it from falling prey to older or newly exploited CVEs. Usage of the Windows operating system, the Microsoft Office suite and its macro functionality is commonplace and necessary in enterprises of all sizes, but the latter is also frequently abused for infecting victims [113], for example through phishing. Using other built-in scripting languages, including VBScript, JScript and PowerShell, which are available natively in Windows, adversaries can execute virtually any type of action.

Limited Business Alignment

Another main obstacle is the lack of alignment between the business and the SOC. The first reason for this is the fact that the SOC often resides in an enclave, which is good in terms of security [106], but can result in limited visibility into changes in the IT environment or its users [7]. Another factor is the fact that the SOC often does not have visibility into the primary business processes performed by an organization, resulting in a lack of contextual information [7, 124, 125, 131]. On the other end of the spectrum are the SOCs that do have visibility into their constituency, but have no idea about what's going on outside of it, missing critical information about developments in the threat landscape [128, 136].

Inadequate Resources

The lack of resources, both in terms of technology as well as in human resources, summarizes the next big category of difficulties we observed. The first technological aspect is the fact that it was not economically feasible to store and retain a large amount of security data in the past [7, 13]. Related to this is the fact that it was nearly impossible to process the amounts of data already being stored [7, 13, 136] and the fact that new, unstructured, data sources were also increasingly being added to conventional databases [13, 132]. As a consequence, many Security Information and Event Management (SIEM) deployments were crippled from the start by not feeding them with the right data to increase visibility [131]. Integrating SIEM technology is hard because they are relatively inflexible [124, 132], many point solutions have to be integrated and they require a lot of knowledge to implement properly. Furthermore, deploying a SOC and its infrastructure is no sinecure: it takes continuous tuning of many knobs in order to make it run effectively [111] and out-of-the-box detection functionality is often insufficient [112]. Analysts are often flooded with alerts that are the result from inadequate default detection method configuration(s) [7], which often also contain numerous false positives [129], resulting in even poorer performance. Security of the SOC infrastructure is another aspect to get right. Reliability and availability of SIEM deployments has been a problem in the past [7, 13].

In addition to the various technological shortcomings, it is also generally well-known that SOCs are having a hard time employing enough people with the right skills. Staff deficiencies have been reported about numerous times by Enterprise Strategy Group (ESG) [130, 131], Enterprise Management Associates (EMA) [124] and the SysAdmin, Audit, Network, and Security (SANS) Institute [137, 138].

Augmented Attacks

The fourth category of difficulties we observe in modern SOCs is related to the number and impact of attacks occurring. The first contributing factor is that adversaries can get access to malicious code relatively easy on (underground) black markets. One example of what can be bought is the so-called exploit kit, which offer adversaries various capabilities to easily infect their targets [62] with ransomware, for example. Poly- and meta-

morphic malware are also commonplace, and packers offer adversaries the capability to avoid detection [3, 66]. Distributed Denial of Service (DDoS) attacks can easily be bought on the same underground marketplaces [108] or even via so-called booters [80], which can be described as commercial services offering "stress-testing" capabilities. Other offerings include botnet rentals, bulletproof hosting and consulting services offering help in setting up various attacks [66]. No reliable statistics are available regarding the size and growth of cybercrime [25], but what is certainly true, is that it has evolved considerably over the years [24, 66].

Adversaries are also increasingly making use of automated attacks targeting low-hanging fruit. Scanning and analyzing the entire Internet IP space has become easily feasible with the emergence of software like ZMap [37] and platforms such as Shodan[4] and Censys [38]. Crawling the web for websites running on outdated software or making use of insecure extensions is performed constantly. Platforms such as WordPress and Magento are under constant scrutiny by attackers employing scanning tools such as WPScan[5], Nikto[6] and Arachni[7] to find exploitable websites and applications. The successful compromise of a website can have various consequences, such as leakage of confidential data, becoming part of a botnet or other attacker infrastructure and serving exploit kits. Sometimes the compromise can lead to the attacker gaining full control over a machine, which can then serve as a stepping stone in a more advanced attack.

Advanced Persistent Threats (APTs) are at the other end of the spectrum. They can be described as adversaries having access to significant resources and expertise [49]. APTs have specific targets and clear objectives, are well-organized, have access to many resources, operate over prolonged periods of time and use stealth and evasion techniques during their attacks [17]. These characteristics define what APTs and distinguish them from more traditional threats. Attacks by APTs are often called campaigns due to their long-term, multi-phase approach, which can be mapped to a construct called the *cyber kill chain* [48]. Various tactics used by APTs along the stages of the kill chain include *spear-phishing*, *social engineering*, *watering hole*

4 https://www.shodan.io/
5 https://wpvulndb.com/
6 https://cirt.net/Nikto2
7 http://www.arachni-scanner.com/

attacks, *data hiding* and the use of so-called un-patched and unknown zero-day vulnerabilities [17]. They may employ custom malware that is not detected by conventional detection methods, such as anti-virus and Intrusion Detection System (IDS), but they will also make use of run-off-the-mill and/or modified malicious code in their attacks. The *Energetic Bear* campaign is an example of this: the APT used "generally available" exploit kits for infecting their targets, but deployed custom backdoors for persistence and remote access [134].

Summary of Challenges

As described in the previous sections, SOCs face numerous difficulties when protecting their constituencies. As illustrated in Figure 4.1, the difficulties that arise are very diverse and can be categorized into four categories. These categories are the increasingly complex IT environment, limited business alignment, inadequate resources with respect to people and technology, and finally, ever-evolving adversaries and attacks. The complexity of the IT environment makes it hard to see and understand what is currently going on within an organization's IT infrastructure. SOCs also often don't know what the most critical systems are and suffer from a limited view of the business processes. Not having enough resources in terms of technology and people make it next to impossible to protect against ever-evolving threats. All of these aspects make it harder for a SOC to adequately protect its constituency.

To address these challenges, vendors offer scalable and performant products that promise to solve many if not all of the difficulties. Many vendors also promise excessive results in prevention and detection rates when deploying their solutions and boast effortless deployment and maintenance. Although technology is certainly a piece of the puzzle for solving the challenges faced by SOCs, it is not as easy as buying a box and deploying it to a data center to operationalize it. We also think there is no single market for *Security Analytics*, but that it consists of several, related, types of solutions, which often have to inter-operate with each other and with older types of solutions. These new developments in the security industry also require more specialist knowledge, in fields like machine learning and threat intelligence for example, to be developed within the SOC.

We conclude there are many aspects SOCs need to consider, and we would like to pose the following as the main research problem:

> Security Operations Centers fail to realize the potential of *Data Driven Security Operations* in order to better protect their constituency.

4.2 PROBLEM JUSTIFICATION

As described before, Security Operations Centers (SOCs) face various difficulties, which we attribute to many different causes as shown in Figure 4.1. SOCs and the organizations they belong to, have always had a need for data in order to prevent, detect and respond to attacks and intrusions. This started with Intrusion Detection Systems (IDSs) and rudimentary log management solutions, the latter of which were eventually superseded by Security Information and Event Management (SIEM) solutions, which are still used as a core technology for log management, alerting and dashboards.

Deploying and maintaining a SIEM is not an easy process however and many deployments have failed at some stage, resulting in not realizing the full potential of a SIEM solution. The data storage and processing capacity of SIEM solutions is also often said to be insufficient, resulting in crippled deployments failing to capture all of the necessary data. In addition to the volume of data, there's also a wide variety of logging formats that have to be captured, in some cases even including full packet captures (PCAPs). Besides that, SIEM solutions demand proper tuning of correlation and detection rules to provide value to security analysts.

Collecting data necessary to perform compliance or security monitoring is only part of the story. When analysts can't get access to the data in a timely manner to perform their analysis, they can't decide which events have to be investigated more closely. To improve analysis, available data should be contextualized by adding business specific meta-data, but many SIEM solutions offer only rudimentary approaches for this functionality. Analysts also can't use the data that is already available to search retroactively for newly released Indicators of Compromise (IOCs) or looking for threats that have passed the detectors

unseen, because currently used technologies were not designed for that purpose.

All of these aspects revolve around data, including many operations performed on data, such as data collection, storage and analysis. Helping SOCs get a better understanding of what it takes to realize a data driven approach to security operations will eventually allow them to respond faster when necessary, improving the risk profile of their constituencies.

5 ARTIFACT REQUIREMENTS

To address our research questions and the problem posed in Chapter 4 we propose to create two distinct artifacts. Both artifacts will have their own requirements, which we describe in the following sections.

5.1 CONCEPTUAL MODEL REQUIREMENTS

The first artifact that we will design is a conceptual model. The conceptual model serves the purpose of showing Security Operations Centers (SOCs) what aspects play a role in successful *Data Driven Security Operations*. We chose a conceptual model as the artifact for this specific problem, because we want to capture several inter-related concepts. A model can give people a single view of *Data Driven Security Operations* and what it consists of, increasing their understanding of the concept. The requirements for the conceptual model are listed in Table 5.1.

Table 5.1: Model requirements.

ID	Requirement
R1	The conceptual model should describe what *Data Driven Security Operations* consists of.
R2	The conceptual model should be generic, i.e. applicable to many (different types of) organizations and their SOCs.
R3	The conceptual model should be complete, i.e. it should address all parts of the concept.
R4	The conceptual model should be coherent, i.e. its parts should fit together consistently and logically.
R5	The conceptual model should be easily comprehensible, i.e. it should make sense without a lot of explanation.

The main functional requirement, R1, is based on our hypothesis that organizations and their SOCs do not have a clear, complete view of what *Data Driven Security Operations* consists of. Requirement R1 is also important to be able to answer the second research question, which requires is about describing the main aspects of *Data Driven Security Operations*. We picked requirements R2 and R3 to increase the usability of the model in practice. The model should be generic in order to be applicable to different types of SOCs and a general audience. SOCs can be different in a wide range of aspects, including market vertical(s) they operate in, the size of the constituency it protects, its own size, how the SOC fits in with the constituency (e.g. does it run their own SOC or does it use the services of a Managed Security Services Provider (MSSP)) and and what capabilities are provided by the SOC. The model should also not leave out important aspects of *Data Driven Security Operations*, which would make it far less usable in practice. Requirements R4 and R5 are in place to make the model understandable, which may result in better results bringing the model into practice, which is, ultimately, the goal of this thesis. These two requirements are also directly related to the third research question of this thesis, because the presentation of the conceptual model is closely related to its coherence and comprehensibility.

5.2 INSTANTIATION REQUIREMENTS

Evaluating the conceptual model on its own would not be sufficient to evaluate and capture its usefulness in practice, which is why we decided to design a second artifact. The second artifact is an implementation of the conceptual model that can be used in practice. In addition to providing a way to interact with the model it also helps evaluating the conceptual model. The implementation has a separate set of requirements, which are listed in Table 5.2.

The research goals of this thesis include making SOCs understand what *Data Driven Security Operations* is, what it consists of and finding out how they can position themselves on each of the aspects of *Data Driven Security Operations*. The first part of the goal is partly realized by the creation and presentation of the conceptual model. We thought of a second set of requirements that cover the practical use of an instantiation of the conceptual model to offer something practical. The prac-

Table 5.2: Instantiation requirements.

ID	Requirement
R6	The instantiation should be usable in practice.
R7	The instantiation should offer a SOC practical value.
R8	The instantiation should offer a SOC the ability to assess its maturity within *Data Driven Security Operations*.
R9	The instantiation should make it possible for SOCs to compare themselves against themselves and others.
R10	The instantiation should show a SOC where it falls short and what it can consider to improve.

ticality of the instantiation is captured in R6: the instantiation should be usable in practice, e.g. there should be a way to interact with it and should fit a SOC's needs. Requirement R7 is closely related, because it depends on R6: if the instantiation is not usable in practice, it would most likely not offer any practical value to the SOC using it. We want the SOC to gain something concrete from the instantiation. Requirement R8 is a result of the third research question of this thesis and forms one of the main functional requirements for the instantiation. The instantiation should allow SOCs to measure themselves on a scale within the model. Requirements R9 and R10 depend on requirement R8. These two requirements increase the practical value that SOCs receive from using the instantiation. By being able to compare, SOCs can see how they are doing compared to other SOCs or they can see how their status is changing over time. Fulfilling R10 offers direct insight into where a SOC falls short with respect to each of the aspects of *Data Driven Security Operations* and should result in concrete pointers for improvement.

6 DESIGN OF ARTIFACTS

In this chapter we describe the design of the two artifacts based on the requirements listed in Chapter 5 We choose to present the final results of the design phase for the conceptual model first and explain how it came to fruition afterwards. The reason for this is that the description of the process may shroud the occasional reader in seemingly unnecessary intermediate steps and results. These were essential for construction of the final artifacts, however, which is true especially for the conceptual model. We then describe the design rationale and implementation of the instantiation.

6.1 CONCEPTUAL MODEL

Our model as shown in Figure 6.1 is the final result from a number of trials to construct a single mental image for our vision of *Data Driven Security Operations*. The model consists of six facets: *Situational Awareness*, *Threat Intelligence*, *Detection Methods*, *Response & Investigation*, *SOC Staff* and *Security Operations Center (SOC) Infrastructure*. At first sight these may seem unrelated, but what interconnects them is the fact that they all revolve around data. We describe this in the following subsections. We will first introduce and summarize each of the six facets, after which we provide a full description of the entire model and how the facets fit together.

6.1.1 Data Driven Security Operations Summarized

Our vision for the Data Driven Security Operations Center is presented as a single conceptual model composed of six strongly inter-related facets. We will now introduce each of the facets and shortly summarize them. A full description of the facets and the model as a whole is described in Section 6.1.3.

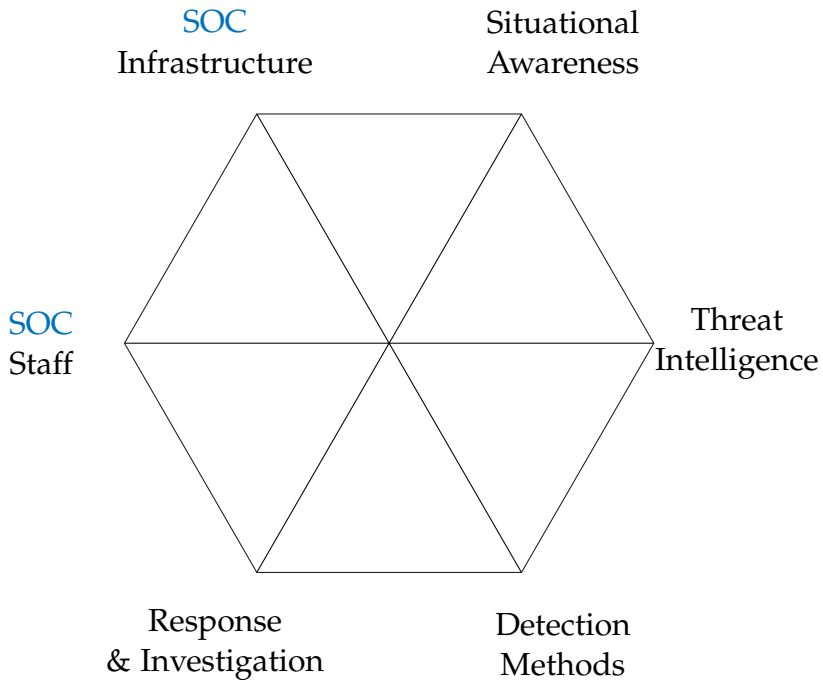


Figure 6.1: Conceptual model for *Data Driven Security Operations*.

Situational Awareness

Defined by Vidulich et al. [100], Situational Awareness is the *continuous extraction of environmental information, integration of this information with previous knowledge to form a coherent mental picture, and the use of that picture in directing further perception and anticipating future events*. In the information security domain it consists of a three-phase process consisting of *situation recognition*, *comprehension* and *projection* [55]. Current approaches to acquiring Situational Awareness include vulnerability analysis through attack graphs, intrusion detection and correlation, attack trend analysis, causality analysis and forensics, taint and information flow analysis, damage assessment and intrusion response [55].

Before being able to perform any of these activities, it is necessary to have the required data available. This includes data with regards to assets (both hardware and software), how these assets are connected together and what is currently happening on and with the network and assets. Concretely, some data sources to be inluded are logs from endpoint hosts (workstations, laptops, servers), applications and services, operating systems, network devices and security devices. In addition to those, contextual types of data can also be added, including configuration and vulnerability management data, user identi-

ties, geo-information, whois data, (passive) dns data and reputation data. Information extraction, integration and analysis from all of the available data then has to be performed in real-time in order to get a view of the current situation [91]. This real-time view has to be constructed at different levels of abstraction and detail, namely on the strategic, operational as well as the tactical level, to support decision making on all of these levels.

Information visualization is the cornerstone for attaining Situational Awareness [60]. As long as machines are not intelligent enough to understand they are under attack, it is up to a human analyst to comprehend and project the current situation. Effective dashboards provide an analyst with an overview of the current situation supplying him with information extracted and enriched with context [75]. Creation of effective dashboards is not trivial, however [56, 67]. A model of the IT environment showing interdependencies between systems augmented with security-relevant details can help provide an analyst with intuitive insights about the IT infrastructure and allows him to quickly respond to new information [61].

Threat Intelligence

The second facet of *Data Driven Security Operations* is Threat Intelligence. Gartner defines it as *evidence-based knowledge, including context, mechanisms, indicators, implications and actionable advice, about an existing or emerging menace or hazard to assets that can be used to inform decisions regarding the subject's response to that menace or hazard* [123]. As with any type of intelligence created, it is aimed at providing information and knowledge that can aid in decision making [34]. Examples of the usage of Threat Intelligence in information security include preventing attacks, decreasing the time to analyze an attack and showing insights about the current risk landscape [20].

We distinguish Threat Intelligence on three distinct levels originating from military doctrine: the strategic, operational and tactical levels, together forming the three *levels of war* [73]. Chismon and Ruks [20] chose to include a technical level, but we see this as a subset of tactical Threat Intelligence, which is also supported by Friedman and Bouchard [115]. Threat Intelligence on the strategic level provides internal as well as external high level information to improve decision making and prioritizing risks. It is typically consumed by members of the board and people reporting to them, such as the Chief Information

Security Officer (CISO), and usually does not contain technical information [20]. On the operational level, Threat Intelligence informs SOCs about current and impending attacks against the constituency and context around these. For a typical SOC acquiring operational Threat Intelligence can be hard because of limited insight into the communication channels and infrastructure used by (potential) adversaries. The SOC then depends on reports distributed by organizations who do have access to this kind of information in order to gain insights. Operational Threat Intelligence is consumed by the higher layers of the security staff, such as the security and SOC managers [20], and can be used to respond to an attack with adequate measures or mitigate an impeding attack by operationalizing the right controls.

The last type of Threat Intelligence, the tactical kind, is aimed at getting to know what adversaries use to attack the constituency. On a high level and long-term view, the SOC members can use this kind of intelligence to increase their knowledge about the usage of specific Tactics, Techniques and Procedures (TTP) that adversaries use [20]. On the short-term and low level view the SOC can collect and integrate Indicators of Compromise (IOCs) with existing security solutions to improve the current protective, detective and responsive controls. One example of this is using tactical Threat Intelligence to (in)validate and prioritize alerts generated by various systems [115]. Due to the amount of low level data, which includes Internet Protocol (IP) addresses, domain names, file hashes and reputation data, for example, the processes to operationalize the data should be highly automated. Several efforts exist to standardize formats for creating and distributing Threat Intelligence, such as Cyber Observable eXpression (CybOX) [5], Structured Threat Information eXpression (STIX) [4], Trusted Automated eXchange of Indicator Information (TAXII) [28] and the Incident Object Description Exchange Format (IODEF) [31], increasing the ability to automate ingestion and production.

What is true for all levels of Threat Intelligence and for a Threat Intelligence program to be successful is that there should be well-defined processes in place that support requirements elicitation, data collection and analysis, evaluation and optionally, production and sharing [20, 115]. Evaluation of the usage of Threat Intelligence is important on all levels, especially on the tactical level, because it's not hard for an attacker to create

new malware or spin up new a new infrastructure that can't be identified by existing signatures, for example.

Detection Methods

Despite numerous investments in security solutions to improve detection rates, adversaries continue being successful breaching organizations. Security Information and Event Management (SIEM) systems, which can be considered one of the most advanced technologies available to SOCs currently, rely on static rule sets manually implemented and maintained by human operators to alert on undesirable behavior. In many cases these rules are not tuned well enough resulting in many false positive alerts being reported that have to be triaged [130]. Intrusion Detection Systems (IDSs) and Intrusion Prevention Systems (IPSs) are a different type of system that can be based on signature-based and anomaly-based detection, both of which can be evaded [39]. The idea of data fusion to increase the effectiveness of IDSs was presented by Bass [6] in 2000 already, which he described as "requiring the integration of numerous diverse disciplines such as statistics, artificial intelligence, signal processing, pattern recognition, cognitive theory, detection theory, and decision theory".

The application of machine learning, statistics and anomaly detection to detect threats has been researched extensively, as can be concluded from the reviews conducted by Tsai et al. [96], Jyothsna, Prasad, and Prasad [58] and Liao et al. [64], but deploying these in operational settings apparently still presents numerous challenges [41, 86]. These include low detection efficiencies, the absence of assessment methodologies, the high cost of analysis [41], the high cost of false positives and the semantic gap between detection results and their operational interpretation [86]. Despite these challenges, we think current methods for detection can be improved by adding additional context to the data that is used to generate alerts on. One of the consequences of this is that an alert can make use of the additional data to provide the analyst with extra context, which might result in a quicker triage of the alert. The application of additional data types, data sources and the fusion of these can improve the results of IDSs [6] and increase the number of features that can be selected for use in machine learning models [107].

What is important whatever method is used to analyze and mine data, is to follow well-defined processes for doing so. The

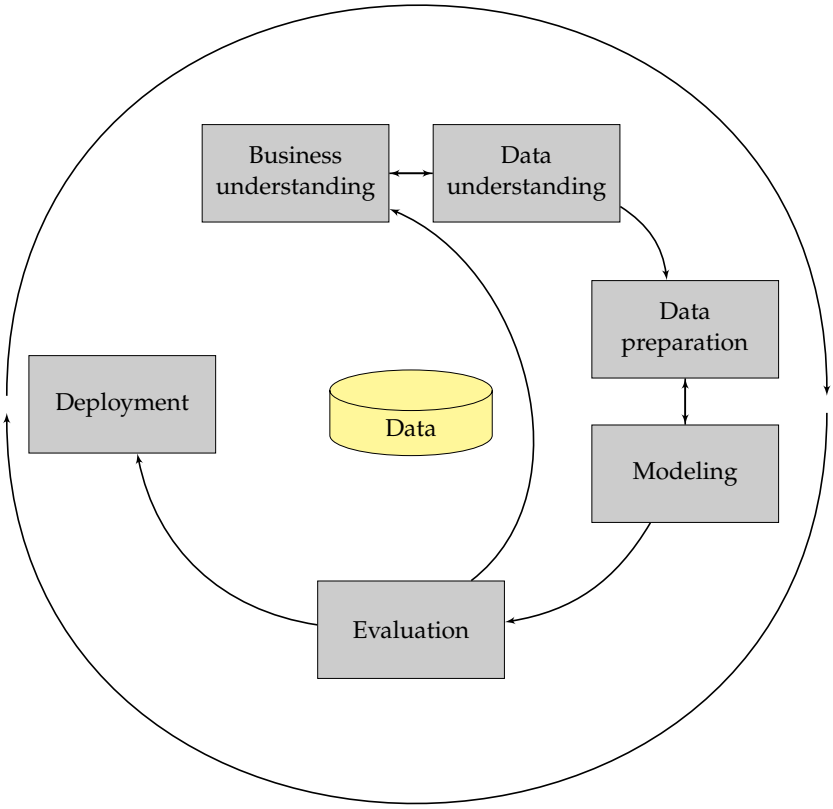


Figure 6.2: The CRISP-DM reference model for data mining.

CRoss-Industry Standard Process for Data Mining (CRISP-DM) [83] reference model for data mining is an example of such a process. The six phases of the model, *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation* and *deployment*, described on a high level, capture the essential steps for performing data mining.

As described by Virvilis and Gritzalis [101], the use of widely accepted security mechanisms such as anti-virus (AV), IDS and IPS can help, but falls short when an organization is targeted by a so-called Advanced Persistent Threat (APT). They propose a highly integrated architecture, where low severity events generated by adversaries (or malware), which are inevitable created by them to reach their goal, are correlated with historical events from a wide range of sources. Changes in the IT environment,

no matter how small, will eventually end up in alerts this way. As surveyed by Akoglu, Tong, and Koutra [2], research into anomaly detection on graph data has been ample and shows clear advantages, such as offering a powerful representation and making use of the relational nature of the problem domain. But as described before, all *Detection Methods* should be small in scope, depend on a clear threat model and be properly evaluated [86].

Response & Investigation

A security incident is a violation or imminent threat of violation of computer security policies, acceptable use policies or standard security practices [59]. What is considered a security incident within a specific organization depends on its business processes, but common events that may trigger an incident include unauthorized system access, malware infections and data loss [105]. It is critical to timely respond to security incidents in a systematic way and to prevent the same incident from happening again in the future [22]. As soon as a security incident is detected, it has to be analyzed to identify its root cause in order to resolve the incident and remediate it. Irrespective of which party is responsible for the incident handling process, having access to the right data is an absolute must within incident response. The data is necessary in order to investigate the incident, e.g. find out what happened, who or what entities were involved, assess the impact of the incident, how to recover from the incident and to actually recover from the incident. During the incident response process every second counts, so having access to the data and the right fidelity of data is crucial. Ideally the data is already available, but when this is not the case, the analyst needs the capability to collect the data ad-hoc, so having access to and control over endpoints at the time of an incident can give unprecedented insight into the IT environment at the time of an incident. Some concrete data sources include log data from various types of devices and applications and contextual data. Memory and packet captures contain high fidelity data and are becoming more important during forensic investigations.

Investigation into incidents is mostly initiated only after an incident response process has been started and when deemed necessary: reactive investigation. We propose its logical counterpart, proactive security investigation (also popularly called *hunting*), to become a core part of the SOCs functions. Proactive

investigation should be used to uncover threats hiding within the security data that is already available but have not been found thus far and for hypothesizing future attacks and incidents. In a mature *data driven security* strategy, we foresee the emergence of the continuous security response process. The continuous security response process consolidates the prevention of, detection of and response to threats and security events and incidents. In a sense it is the culmination of the *Data Driven Security Operations* strategy, in which data is transformed into information, then intelligence and eventually applied in practice to improve all of the SOC functions. Well-defined and repeatable processes for data collection, analysis and incident remediation are the main tenets in Response & Investigation. Jacobs and Rudis also describe a process for performing data driven security in [54, ch. 12] that does not have to be applied in reactive security investigations, but can also be used to increase *Situational Awareness* or developing new *Detection Methods*. Other scenarios for applying analytics in information security are described by Talabis et al. [93].

SOC Staff

The fifth facet of *Data Driven Security Operations* adds the human element to the model: the SOC staff. Hiring and retaining the right people in information security can be problematic because skilled people are scarce [109]. As described in Section 4.1, this is also a problem that SOCs are experiencing. SOCs should be looking for people having a diverse skill set, the right mind set and having a (practical) background in IT [106].

The *Data Driven* SOC needs a diverse skill set, just like a 'normal' SOC. This includes both theoretical and practical knowledge about network devices, network security analysis, network protocols, application security and engineering. Operators need to have experience in using the software and tools the SOC employs to deliver services to its constituency, including the SIEM and Network Security Monitoring (NSM) solutions. Depending on the focus of and capabilities offered by the SOC, it needs skills related to reverse engineering, malware analysis and digital forensics.

To become *data driven* the SOC needs skills in additional fields, all depending on what services are and will be provided by the SOC and which capabilities the SOC wants to improve. For example, when the SOC wants to improve the current detection rates or decrease false positives of the SIEM or IDSs by

employing machine learning or statistical methods, expertise in these areas is necessary. This is also exemplified by Sommer and Paxson [86], who state that understanding what a machine learning model is doing, evaluating models, reducing the scope of models and understanding what the limitations of the machine learning models are, are critical for successfully deploying machine learning in an information security setting.

When the SOC wants to improve its usage of Threat Intelligence it needs different skills at each of the levels, e.g. programming and engineering for integrating and automating tactical Threat Intelligence [106] and analytical and communication skills at the operational and strategic levels to improve intelligence dissemination. A prime example of improving and formalizing human reasoning about threats (and intrusions) is the Diamond Model by Caltagirone, Pendergast, and Betz [10]. Another example is the concept of a Kill Chain as presented by Hutchins, Cloppert, and Amin [48] which can be used to reason about the steps an adversary has performed during an attack.

When the SOC wants to start or improve its proactive investigation and response capabilities, it may need skills in machine learning, statistics, engineering and visualization, depending on its focus. All of the capabilities offered by the SOC are supported by technology, and because of the increasing scale of data collection, processing and analysis there is a need for practical knowledge about implementing and maintaining different types of systems than were used in the past, such as distributed data storage and processing frameworks.

The threat landscape is constantly evolving, demanding information security professionals to remain up-to-date. The SOC should encourage its operators to continuously learn about what is currently going on in the field and keep their knowledge up to par. SOCs should have programs in place for professional certifications and should reinforce self-education.

The SOC should also have predefined workflows and methods for collaboration in place that support the various capabilities offered. Examples of these include protocols for alert triage, incident escalation, incident hand-off and incident response. These help operators to effectively and efficiently do their job and foster repeatability, consistency and communication. Care has to be taken to not overimpose the operators with too much structure, however, as this may negatively impact the SOC's effectiveness [106].

The final aspect we think is critical is defining the right metrics for measuring SOC effectiveness and success. Technically speaking this is not only about SOC staff, but because of the fact that the actual work in a SOC is still done by human operators, makes this the right place to introduce them. "Security is a process" [81], or actually a set of processes, that need to be measured, for which one needs metrics and key indicators. Jaquith [56] defines the following characteristics for good metrics:

- Consistently measured
- Cheap to gather
- Expressed as a cardinal number or percentage
- Expressed using at least one unit of measure
- Contextually specific

In our opinion SOCs should have metrics in place for most, if not all of the capabilities offered. Some general metrics with respect to detection and response (*resolution*) have been identified by Mandiant [121]. In their whitepaper they define metrics for the time to *Detect, Review, Analyze, Identify, Notify, Collect, Validate and React* to incidents. We think these metrics are a good start for measuring general effectiveness of the SOC at a high level, but should be extended with both quantitative and qualitative metrics aimed at measuring lower level and specific aspects of operating the SOC, such as the usage of Threat Intelligence, effectiveness of Detection Methods and the level of Situational Awareness realized.

SOC Infrastructure

The final facet of *Data Driven Security Operations* is *SOC Infrastructure*. All of the capabilities offered by the SOC, such as security monitoring, threat analysis, and incident response, should be supported by the SOC infrastructure. The SOC infrastructure should be seen as the complete set of technologies used to enable the SOC to execute its mission. It includes the systems running on network devices and endpoints to report events (agents), the (central) data repositories, the network through which the data is transmitted and systems making use of the data that is stored. Because of the fact that no IT environment

is the same any security technology should be flexible, extensible and compatible in terms of data integration from heterogeneous sources [107]. The SOC infrastructure should be capable of large scale data storage, processing, analysis and visualization with high performance [130]. Because of the SOC's high reliance on the infrastructure, care should be taken that the systems they employ stay available at all times, actions are audited and data is protected from being altered. Several research directions in security for *big data* have been identified by Rajan, Ginkel, and Sundaresan [133], such as *granular access control*, best practices for non-relational data stores and data provenance.

6.1.2 Formalizing the Conceptual Model

As we described before, the model as shown in Figure 6.1 and explained in the previous section and its subsections is the result of several trials to create a single mental image for our vision of what *Data Driven Security Operations* consists of. The six facets, *Situational Awareness*, *Threat Intelligence*, *Detection Methods*, *Response & Investigation*, *SOC Staff* and *SOC Infrastructure*, have been described and substantiated with (scientific) literature and existing surveys from the industry. We chose these six facets because in our eyes they fit together and form a coherent and comprehensible whole, which we reason about in the next part.

Situational Awareness, the first facet in our model, has actually already been studied extensively, as indicated by the review conducted by Franke and Brynielsson [40]. In their review they analyzed 121 relevant articles, surfacing a diverse set of directions within Situational Awareness research. These include specialized application fields, designs and implementations for improving situational awareness and information fusing. In our opinion, attaining *Situational Awareness* is an important requirement for SOCs to realize their potential, because without a true understanding of what is going on within the IT environment, the SOC will not be able to adequately respond to attacks and security incidents. In order to attain *Situational Awareness* the SOC needs to collect, aggregate, store, process, integrate, visualize and analyze various types of data for a human operator to make sense out of it.

We contemplated about adding *Threat Intelligence* to the *Situational Awareness* facet, because that would fit with the mili-

tary notion of intelligence: "it provides the commander a variety of assessments and estimates that facilitate understanding the OE", "allows anticipation or prediction of future situations and circumstances and it informs decisions by illuminating the differences in available COAs" [34]. The Operational Environment (OE) for the SOC could then be defined as the IT environment of its constituency located within the threat landscape surrounding it. *Threat Intelligence* could then be seen as intelligence informing the SOC about what is going on in the threat landscape, which it can use to increase *Situational Awareness*. Another indicator that *Threat Intelligence* can be seen as part of the *Situational Awareness* facet is the fact that we used the levels of war in both facets to indicate the level of abstraction and detail with which the IT environment and *Threat Intelligence* are presented. The reason for introducing *Threat Intelligence* as a facet is because of practical reasons: the *Situational Awareness* facet would otherwise have become too large to address appropriately.

The next two facets, *Detection Methods* and *Reponse & Investigation* are about the (broad, high level) operations performed by the SOC. At one side there's the need to detect an attack or compromise as soon as possible. The SOC has various types of technology available to realize this, such as firewalls, IDSs and one or more SIEM instances. In many cases these systems are not tuned well enough resulting in many false positive alerts being reported that have to be triaged. Tuning these systems is a solution, but maintaining the correlation rules of a SIEM deployment is no sinecure and they are often configured to result in no false negatives, letting false positives through. Despite the challenges involved we think that the field of machine learning has improved in ways that make it possible to better support the challenges present in information security. The addition of contextual and other heterogeneous sources of data for analysis can result in novel features to be used in machine learning models or providing analysts with additional information for quicker, manual triage. On the other hand of the spectrum is *Response & Investigation*, for which the SOC needs access to relevant sources of data to perform security investigations. In the ideal case the SOC can readily access the necessary data from a central data repository or has means to obtain it remotely. Having access to the forensic data also allows a SOC to pro-actively investigate the IT environment, possibly uncovering adversaries that went undetected. After having access to

the data, a way to improve the efficiency of a security investigation is linking data originating from different sources together so that an analyst does not have to do this manually.

The final two facets, *SOC Staff* and *SOC Infrastructure*, finish off the conceptual model for *Data Driven Security Operations*. Without the right people that actually have the knowledge and skills to perform security monitoring, incident response and all of the task associated with those, the SOC will not be effective. In the *data driven* SOC there's not only a need for the usual skills required for security monitoring and incident response, such as knowledge about networking, application security and security analysis, but also in fields like machine learning, statistics and intelligence analysis, for example. The human element within the SOC should also continuously be developed and improved, for example through training and certification and the deployment of standardized workflows. Adequate metrics should be in place to measure the effectiveness of the SOC and to determine where it falls short. The other supporting pillar of the *data driven* SOC is the *SOC Infrastructure* facet. Most of the capabilities offered by the SOC rely on technology, so care must be taken that the right technology is in place. In the *data driven* SOC large amounts of data need to be collected, stored, processed, analyzed and visualized, which require scalable and high performance technologies to be in place. There's also a need for flexibility, extensibility and compatibility in terms of data integration and access to allow the SOC to mold the technology to its specific needs based on the services offered.

So far we have described all of the six facets relatively loosely coupled, but in an actual SOC the connections among the facets are much stronger. The effectiveness of *Detection Methods* does not only depend on what data is available and how this is analyzed, but also on how well *Threat Intelligence* is being integrated for example. *Response & Investigation* can benefit greatly from having a collection of historical IOCs available to retroactively scan the IT environment for the existence of entities related to the IOCs in the IT environment. New *Detection Methods* can be created or improved through performing and the results of *Response & Investigation*. Increasing visibility into the IT environment and thus improving *Situational Awareness* is only possible when the *SOC infrastructure* supports expanding the amount of data that is collected, integrated and visualized. Deployment and maintenance of the *SOC Infrastructure* and integration with the IT environment at scale requires spe-

cific skillsets. Creating an effective *Threat Intelligence* program or developing novel or improving existing *Detection Methods* requires the right knowledge and skills. The aforementioned connections between facets illustrate that becoming a Data Driven Security Operations Center is a complex process requiring a lot of effort and a holistic approach to all of the facets is necessary.

6.1.3 Description of Data Driven Security Operations

A less formal description of the six facets of *Data Driven Security Operations* based on the formal conceptual model was written for consumption by industry peers. This version is available in Appendix B. A gist of the model is available in Appendix C.

6.2 INSTANTIATION

The second artifact created as part of this research is an instantiation of the conceptual model for *Data Driven Security Operations*. We chose to create another artifact because of several reasons, the first of which is the fact that we needed a way to evaluate the requirements we had set for the conceptual model itself. The implementation would complement the semi-structured interviews that we had setup for evaluation of the model. The second reason is that the conceptual model in itself mostly contains theoretical foundations, which could certainly help a Security Operations Centers (SOCs) in establishing a more data driven strategy to security operations, but did not offer practical value directly. We thought that making the SOC think about the model on the basis of predefined questions would offer greater insights into the conceptual model, thereby making it more usable in practice.

As listed in Table 5.2 we wanted our instantiation to be usable in practice, which meant we needed to construct a way for SOCs to interact with the model. The SOC should also be able to set goals, compare itself to other SOCs and itself and the instantiation should provide SOCs practical value. We chose to implement an assessment consisting of questions that we deemed illustrative for the model. The questions were all created using the well known construct of Likert scales, thereby partly addressing requirement R6. Results from the assessment would then be used to create a *radar chart*, showing the SOC where

they fit within the conceptual model, addressing requirements R8, R9 and R10. The final report created by the instantiation would not only embed the radar chart, but also offer the SOC some practical considerations for improvement and reaching its goals, which addresses requirements R7 and R10.

The next sections describe our design rationale and how we implemented the instantiation, including the motivation for the questions and technical details.

6.2.1 Design Rationale

First we had to determine what the instantiation should encompass for a SOC to be usable in practice. We settled fairly quickly on an assessment because this seemed most practical to use in this setting. We also determined that all questions were to become Likert items, because these would allow for basic arithmetic to be used on them and to combine several questions in a single indicator in several ways. All of the Likert items were scaled on a range from 1 to 5, inclusive, using different scales, ranging from the worst to the best possible rating. The questions have not been evenly divided over all facets, because of the difference in breadth the facets encompass. The breadth of the facets is also the reason that the questions we devised are broad in nature, but we have take care to embed the most important aspects from each facet in them.

The main goal of the instantiation was to offer the SOC the ability to assess its maturity within the model for *Data Driven Security Operations* (R8). A radar chart seemed a good fit to realize this, because it is possible to visualize several quantitative variables in a single plot. As described before, the assessment was constructed from several Likert scales, six in total, each belonging to a single facet and consisting of several Likert items. In our instantiation we calculated the mean for each of the facets as usual, i.e. by summing the scores for each of the questions and dividing by the number of questions for a specific facet. Plotting the scores for each of the facets in a single radar chart would then show the SOC how well it scores within our *Data Driven Security Operations* model.

We also wanted to include a way to show a SOC where it falls short and should consider improving (R10). Because SOCs differ in the capabilities they offer, we introduced the setting of a goal level for each of the facets, such that SOCs can set goals for individual facets. Because in our opinion a SOC should have

some leeway in assessing its score for specific facets, we also introduced a weighing factor. Using this factor a SOC can prioritize certain facets over others, resulting in easier or harder to reach goals. The final part of the design of the assessment would gauge the *perceived level*. This perceived level would help us to determine the validity of the questions posed in the assessment: when the perceived level of a facet is close to the score based on the answers, that's an indication the questions posed in the assessment fit their purpose of assessing a SOC's maturity for that facet within the model for *Data Driven Security Operations*.

6.2.2 Implementation

Our implementation consists of a front end and a back end that together result in a working instantiation for assessing a SOC's maturity within our conceptual model. We will describe these parts in the next sections.

Assessment

Determining what questions to ask was the most important part of creating the instantiation. Because of the broad nature of both the conceptual model and each of the facets, we had to either include many detailed questions, resulting in a long test, or relatively high level and broad questions, resulting in a shorter test. Because this is the first version of the implementation, we decided to create a shorter test, with questions broad in nature, but still touching upon each of the important aspects of the facets. This way the questions would engage participants to think about and comment on them, increasing the amount of qualitative data to assess. We had already decided to include only Likert items in the assessment because this would allow us to combine several questions into a single measure. Table D.1 lists the five Likert ratings that we applied. All of the questions posed in the assessment are listed in the tables in Appendix D. The last column shows the scale that was used for a specific question corresponding to Table D.1.

All of the questions were implemented in a Google Forms form, the front end of the instantiation. Google Forms was chosen as the front end because it supports all types of questions necessary, especially the Likert scale. In addition to that, Google Forms can be accessed via an Application Program-

ming Interface (API), through which the data entries can be accessed directly, and allows exporting the data to comma-separated value (CSV) format, which can be read in many software packages and is practical for data processing. The elements of the form were constructed in the following order:

- Identifier for participant
- Weights for facets: six times a value from $\{1, 2, 3, 4, 5\}$
- Questions for Situational Awareness (Table D.2)
- Questions for Threat Intelligence (Table D.3)
- Questions for Detection Methods (Table D.4)
- Questions for Response & Investigation (Table D.5)
- Questions for SOC Staff (Table D.6)
- Questions for SOC Infrastructure (Table D.7)
- Perceived level for each of the facets: six times a value from $\{1, 2, 3, 4, 5\}$

All of the questions in Appendix D were added to the form in the same order as they appear in Tables D.2 to D.7. For each facet two additional questions were added at the end of each part: a mandatory goal level and an optional field for adding notes and remarks.

Report Generator

We implemented a back end in Python for processing the submitted answers. It connects to Google Forms via its API to retrieve the answers to the questions, but can also read them from a CSV file. By default it always picks the most recent submission for a certain organization *identifier* submitted by the participant. It performs several checks to determine the validity of an entry, after which it starts the actual processing. The capped weighted average score for all of the n questions belonging to each of the six ($i \in \{1..6\}$) facets f_i having weight w_{f_i} is calculated as follows:

$$\max(\min((-(w_{f_i} - 3.0) \cdot 0.1 + 1) \cdot \frac{1}{n_{f_i}} \sum_{j=1}^{n_{f_i}} f_{i_j}, 5.0), 1.0) \quad (6.1)$$

These capped average values are then used together with the goal levels for constructing the radar chart data to be embedded in the report.

In addition to showing the radar chart, the report also contains a management summary of the conclusions that can be drawn from the radar chart. These are short descriptions of what is visualized in the chart. The report also contains a full description of the six facets, which we included in Appendix B.

The final part of the report contains concrete recommendations for a SOC based on the answers submitted during the assessment. These recommendations are linked to groupings of the questions, which are indicated by the Roman numerals in Tables D.2 to D.7 in Appendix D. These groupings are listed per facet: the Roman numeral I in Table D.2 for Situational Awareness does not correspond to the Roman numeral I in Table D.3 for Threat Intelligence, for example. The recommendations include specific products, literature and advice, resulting from the high-level market analysis and literature research performed. By default the SOC will only receive recommendations for the average score of the groupings of questions within the facets that score lower than the goal set by the SOC. The order of appearance of recommendations depends on the absolute difference between the average score of the grouping and the goal level, in order of highest to lowest difference. In the report we stress the fact that we are not affiliated with any product, both proprietary as well as open source, and point out that these serve illustrative purposes only. Many of the concepts behind the solutions that we recommend are important drivers for the Data Driven Security Operations Center. The complete set of recommendations available within the report generator can be found in Appendix E.

The final report is generated as a single, responsive HyperText Markup Language (HTML) page and contains the radar chart, management summary, descriptions of the facets for *Data Driven Security Operations* and recommendations for the SOC.

So far we haven't discussed how the perceived level that a SOC submits at the end of the test is used. These values are used to create a different radar chart showing the SOC's perceived level and the capped weighted average score that was calculated. This chart is part of the evaluation of the instantiation, which is discussed in Chapter 7.

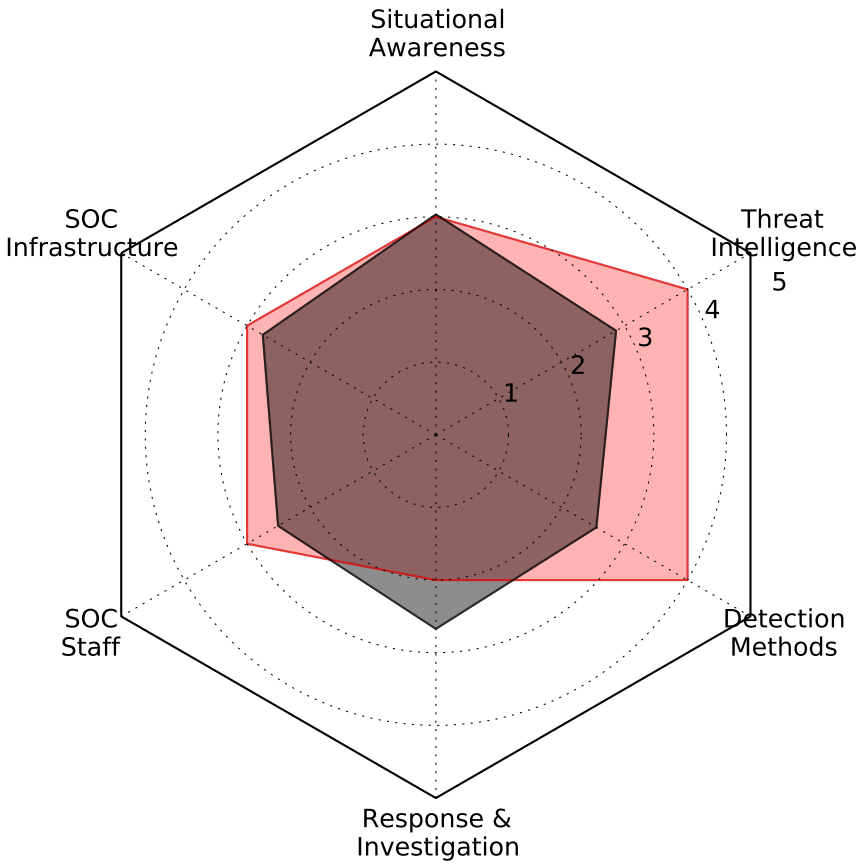


Figure 6.3: An example radar chart. The chart shows the SOC's goal and calculated score in red and black, respectively. The illustration differs from the one in the final report because of a different type of renderer being used.

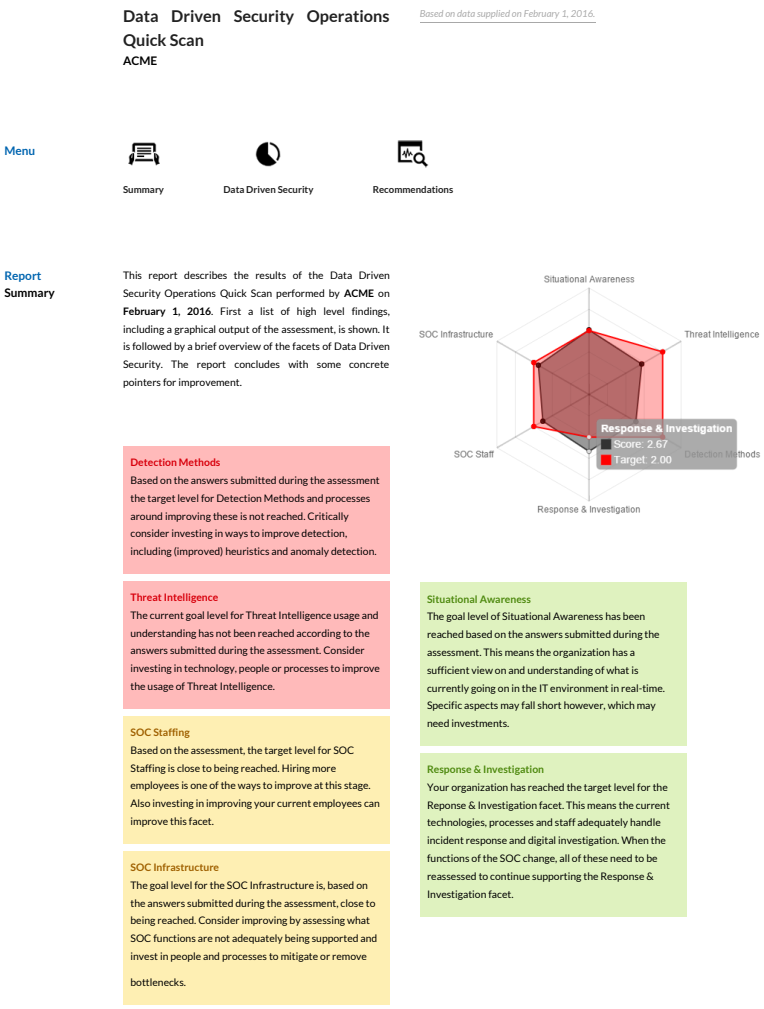


Figure 6.4: Example of the management summary part of a report rendered using a web browser. In a browser the radar chart is created dynamically using JavaScript and incorporates an item legend that appears when hovering over a facet. The colored blocks of text summarize the radar chart. A complete example report can be found on https://hermanslatman.nl/ddsoqs/.

7 EVALUATION

During our research we created both a conceptual model and an instantiation. Both of these were evaluated based on the requirements that were set in advance and listed in Chapter 5. Semi-structured interviews were performed in order to evaluate both the conceptual model as well as its instantiation. We describe the setup of these interviews in Section 7.1. The analysis of the interviews provided us with the necessary data to validate the constructed artifacts, which we describe in Section 7.2. We performed an evaluation of our research strategy in Section 7.3 and the chapter is finished with a discussion of the evaluation in Section 7.4.

7.1 INTERVIEW SETUP

We performed each of the interviews according to a number of predefined steps that are described next. The first part of every evaluation was aimed at getting to know the organization and the interviewee(s). This allowed us to get a view of the Security Operations Center (SOC) and understand how they operate. Getting to know the interviewee was a critical part, because the evaluation of the artifacts relies on the experience and knowledge of the interviewees.

After the introductory part the conceptual model was explained with a short presentation. This included showing the graphical version of the model to the interviewee, explaining what each of the facets means and how it contributes to the model and finally. During the presentation interviewees had the opportunity to ask questions and comment on the model. The questions and comments were answered and have been minuted.

In the third part the interviewees used the instantiation to perform the assessment. They used the implementation of the conceptual model themselves while notes were taken by the interviewer. These notes were taken in order to capture possible misunderstandings and missing information. Any feed-

back received during the walk-through of the assessment was considered very valuable because this could be used for further improvements to the implementation and model.

The final step of each evaluation was a semi-structured interview. It was used to get clear answers with regards to the requirements for the conceptual model and the implementation. Specific questions were asked to establish whether the interviewee understood the conceptual model and to discuss potential shortcomings of the model. The interviewees were also asked to comment on the usability and usefulness of the implementation in practice.

All of the participating interviewees had at least ten years of experience in information security. Most of them have a managerial role in a SOC or are at least leading a team of security specialists within their company. Four organizations were willing to participate within the available time frame, of which two are pure Managed Security Services Providers (MSSPs), one is more akin to a Computer Emergency Response Team (CERT) (but also performing monitoring) and one is a distributor of security solutions who are also increasingly offering managed security services. Several more organizations were considered for cooperation and have been contacted, but these did not respond within the time constraints of this research or were not interested in cooperating.

All of the interviews were transcribed the results of which are available in Appendix A. The participating organizations have reviewed the transcriptions and any remarks with regard to erroneous parts have been resolved. Identifying information has been replaced with generic identifiers in order to prevent identification and potential loss of competitiveness. The information about the participating organizations and the interviews is summarized in Table 7.1.

7.2 INTERVIEW RESULTS

We analyzed the interview transcriptions using the ATLAS.ti 7 software for qualitative data analysis. All of the transcriptions were added as Primary Documents (PDs) to a single Hermeneutical Unit (HU). These were then coded using the requirements posed in Chapter 5 as codes. The results of this process are described in the following subsections.

Table 7.1: Interview participants

Organication	Date	Part.	(Main) Role
DIST	October 15, 2015 & January 18, 2016	A	Lead security analyst
MSSP1	January 6, 2016 & January 21, 2016	B	CEO
CERT	February 26, 2016	C	CERT manager
MSSP2	March 8, 2016	D	SOC manager
MSSP2		E	SOC product manager

Conceptual Model

The first code corresponds to R1 in Table 5.1: *The conceptual model should describe what Data Driven Security Operations consists of*. During each of the interviews we gave a short presentation about what the model entails and what each of the facets means during which there was room for discussion and remarks. We considered this to be part of the description of the conceptual model, because the contents of the presentation are part of what the participants rely on when discussing the model and their understanding of it. We conclude that all participants found the model, including its six facets, to be a clear depiction of what *Data Driven Security Operations* consists of, based on their understanding of the concept, which had been formed through the presentation and discussion. Explaining the facets *Situational Awareness* and *Threat Intelligence*, the first of which is quite abstract and the second broader than some participants (DIST & MSSP2) thought, did have a positive effect on the understanding of these two facets. MSSP2 indicated that the model as presented forms a basis that shows what is important to take into account when operating a Security Operations Center (SOC).

Three out of four participating organizations indicated that the model does not entirely fit with what they do, resulting in a less generic model, corresponding to R2: *The conceptual model should be generic*. Participant B mentioned that *Response & Investigation* are not entirely part of the services they offer, because MSSP1 focuses on security monitoring. The same holds for Participant D, who adds that as an Managed Security Services

Provider (MSSP) you have less control over what you can accomplish at a customer. Participant A indicated that there are differences between large enterprises and small and medium-sized enterprises (SMEs) which affect the facets. Nevertheless, Participant A thinks the model is applicable to various organizations because of the high level on which the model is described resulting in applicability on different levels.

None of the participants thought something was missing from the model, indicating positive evidence for requirement R3: *The conceptual model should be complete*. Participant C indicates that each of the six facets are important, although he thinks the data sharing aspect of *Threat Intelligence* is underexposed. According to Participant C the *SOC Staff* facet did seem outlandish (R4: *The conceptual model should be coherent*), but adding it into the model does make it more complete (R3). Participant Aalso asked why *SOC Staff* is part of the model. He already thought it was an important facet, but wanted to hear why it was included in the model anyway. Participant A also indicated that he didn't see any insufficiencies in the facets, strengthening requirement R4. Participant B indicated that he found the model to be complete (R3) and coherent (R4). As described before, Participant D thought the model forms a base for understanding what are the things to consider when running a SOC. The co-occurrence matrix produced by ATLAS.ti also indicates that statements supporting or opposing the requirements R3 and R4 occur together sometimes, indicating a relationship between the two.

The final requirement, *The conceptual model should be easily comprehensible (R5)*, is generally supported, but with some remarks. When presenting the model to Participant A, he posed several questions with regards to the *Threat Intelligence*, *SOC Staff* and *SOC Infrastructure* facets. Participant D had a general view of the model after he was shown the graphical depiction of the model and had a feeling for understanding what each facet would mean. He did indicate that showing the graphical depiction would not be enough to get a shared understanding of what the model as a whole and what each facet means. In this case the presentation was necessary for him to get a complete view of each of the facets. The same was true for Participant C and Participant B, who indicated that they understood the model and had no further questions after the presentation of the model and the questions that were answered during that presentation.

Instantiation

Quotations for Requirement R6, *The instantiation should be usable in practice*, were most frequent in ATLAS.ti when compared to other requirements. All of the organizations indicated that the instantiation can be used in practice, but do add several remarks to that statement. In addition to using all of the requirements as a code for the analysis, we also added a code *Questions* that was used to mark remarks with regard to (specific) questions and a code *Extra* that was used to indicate interesting tidbits of information, but not necessarily contributing to our research. The co-occurrence table (Table 7.2) shows that many of the quotations that occur for R6 also bear the *Questions* code, indicating a strong relationship between the two. All of the organizations indicated that several questions exist in the assessment that are hard to answer, are not applicable to their services or are too broad. Examples from Table D.2 and Table D.3 include *SA6* (not entirely clear what a threat model is exactly), *TI11* (creation and sharing of Threat Intelligence (TI) are different things) and *SA15* (as a MSSP you don't always have the ability to increase the amount of data you can collect).

According to Participant C the questions are in fact usable as posed now, because they make you think a little longer about them, although some need more explanation. Participant A mentions that the assessment could be performed by several people within the same organization to increase the effectiveness of the questions. He also indicates that the absence of quantitative questions makes it harder to ascertain certain measures, and that the current implementation would currently be better suited for larger organizations. Participant B indicated that the final report directly shows what matters and is intuitive, but that redoing the assessment could result in different answers.

Participant C concludes the radar chart in the report offers insight into where an organization stands, clearly indicates what a SOC should be working on and that the questions make this more clear. The report as presented in its current form certainly offers organizations a practical value (R7: *The instantiation should offer a SOC practical value*), according to Participant A, but for measuring the maturity of an organization more quantitative questions may be necessary (R8: *The instantiation should offer a SOC the ability to assess its maturity within Data Driven Security Operations*). Participant D indicates that performing the as-

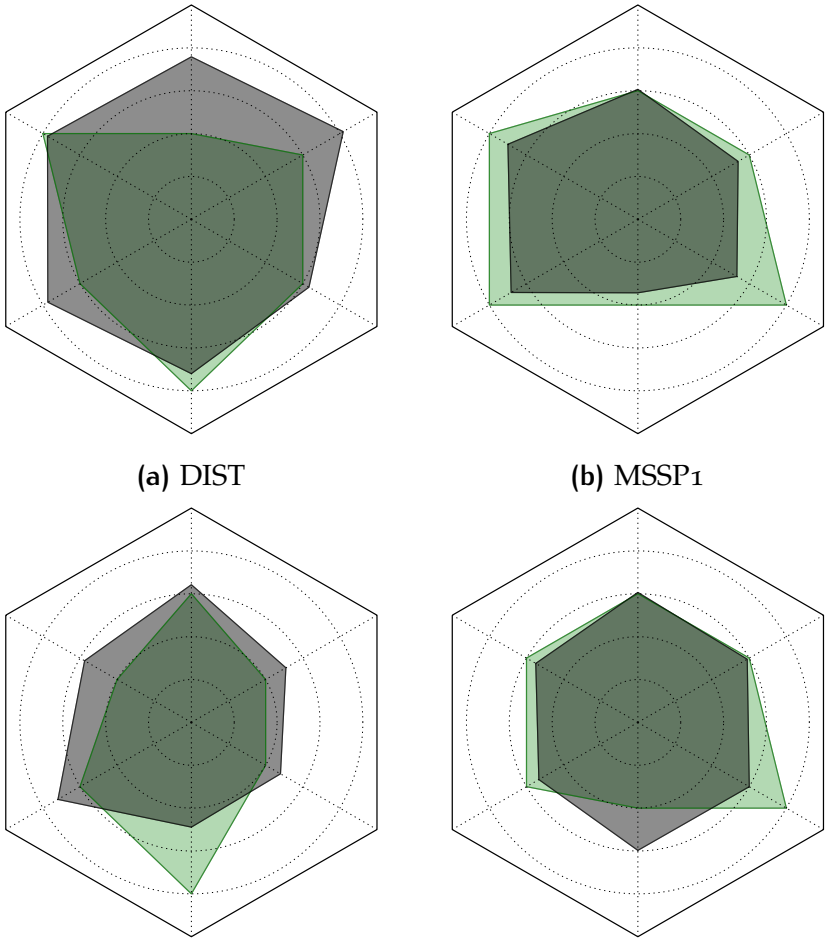
Table 7.2: Code co-occurrence table showing which codes occurred close to each other. R1 to R10 correspond to the requirements in Chapter 5. Additional codes were introduced for pointing to remarks about questions (*Questions* code) and items not directly relevant for our research (*Extra* code). Values are calculated as follows: $c_{ij} = \frac{n_{ij}}{n_i + n_j - n_{ij}}$, where c_{ij} indicates the *c-index* for row i, column j and n indicates the number of occurrences for (a combination) of codes. All *0.00* values, indicating no co-occurrence, have been replaced by a dot.

	Ex.	Qs.	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
Ex.	.	.	.	.	.	.	.	.	.	.	.	.
Qs.	.	.	.	0.03	.	.	.	0.25	0.04	.	.	.
R1	.	.	.	.	.	0.11	0.24	.	.	.	.	.
R2	.	0.03	.	.	0.10	0.06	.	0.02	.	.	.	.
R3	.	.	.	0.10	.	0.15	.	.	.	.	.	.
R4	.		0.11	0.06	0.15	.	0.11	.	.	.	.	.
R5	.	.	0.24	.	.	0.11	.	0.02	.	.	.	.
R6	.	0.25	.	0.02	.	.	0.02	.	0.04	0.02	0.15	.
R7	.	0.04	.	.	.	.	.	0.04	.	.	.	.
R8	.	.	.	.	.	.	.	0.02	.	.	0.08	.
R9	.	.	.	.	.	.	.	0.15	.	0.08	.	.
R10	.	.	.	.	.	.	.	.	.	.	.	.
Tot.	.	0.32	0.35	0.21	0.25	0.43	0.37	0.50	0.08	0.10	0.23	.

sessment increases the understanding of the conceptual model, but that the questions really have to be improved to make it stronger. Both employees from MSSP2liked the idea of adding specific recommendations to the report, increasing support for requirement R7 and R10: *The instantiation should show a SOC where it falls short and what it can consider to improve.*.

Participant A, Participant B and Participant C propose a table with all of the answers to the questions listed to increase the usability (R6), reproducibility (R9) and repeatability (R8). This would improve the quality of comparison of the organization with itself over different periods, resulting in better comparisons. Despite this, Participant C thinks the current implementation can already be used to compare their current status with other organizations and also shows where an organization currently falls short (R10).

Figure 7.1 shows the results of the assessment for each organization in black compared to the perceived level. The perceived level had to be filled in at the end of the assessment by the participant for each facet. Both MSSPs determined their perceived level to be close to the score calculated from their answers, with MSSP2 showing a slight discrepancy on the *Response & Investigation* and *Detection Methods* facets. Participant D admitted he had deliberately scaled his expectation up for the latter. We also see that they both perceived the *Detection Method* facet as better than scored during the assessment, indicating a mismatch between our view of "more mature" *Detection Methods* and theirs. For CERT most facets do not entirely line up with the perceived level, especially for the *Response & Investigation* facet. The biggest discrepancy between the perceived and scored level is observed for DIST, indicating that the questions are suited worst for this organization. We also observe that the scored levels for DIST are relatively high when compared to both MSSPs. As a small distributor of security solutions and security services provider, DIST can be considered a bit outlandish, strengthening the conclusion that the model and instantiation are not entirely applicable to this type of organization.



(a) DIST **(b)** MSSP1

(c) CERT **(d)** MSSP2

Figure 7.1: Perceived level in green compared to the calculated scores for each organization in black. Facet labels have been omitted for readability. Starting from the top and continuing clock-wise, they are ordered as follows: *Situational Awareness*, *Threat Intelligence*, *Detection Methods*, *Response & Investigation*, SOC *Staff* and SOC *Infrastructure*.

Table 7.3: DSR guidelines by Hevner et al. [47]. Shown again for reference.

Guideline	Description
Design as an Artifact	DSR must produce a viable artifact in the form of a construct, a model, a method, or an instantiation.
Problem Relevance	The objective of DSR is to develop technology-based solutions to important and relevant business problems.
Design Evaluation	The utility, quality, and efficacy of a design artifact must be rigorously demonstrated via well-executed evaluation methods.
Research Contributions	Effective DSR must provide clear and verifiable contributions in the areas of the design artifact, design foundations, and/or design methodologies.
Research Rigor	DSR relies upon the application of rigorous methods in both the construction and evaluation of the design artifact.
Design as a Search Process	The search for an effective artifact requires utilizing available means to reach desired ends while satisfying laws in the problem environment
Communication of Research	DSR must be presented effectively both to technology-oriented as well as management-oriented audiences.

7.3 EVALUATION OF THE RESEARCH STRATEGY

As indicated by Hevner et al. [47], usage of the guidelines is not mandatory, but the evaluation of the strategy should be addressed in some manner. We included the guidelines explicitly to strengthen not only the need for this research, but also its execution, evaluation and final conclusions. Table 7.3 shows each of the guidelines by Hevner et al. [47] again as a reference and the argumentations for each of them are described in the following sections.

Design as an Artifact

Our work comprises the design of two distinct artifacts. The conceptual model was created to convey our vision of *Data Driven Security Operations* within the Security Operations Centers (SOCs). The model details six facets that contribute to a full understanding of what is important to consider in a *data driven* SOC. These facets are based on market research, literature, existing surveys and conference materials, such as presentations. Care has been taken to create a generic as well as coherent model out of all of the material that was available. The second artifact is an instantiation of the conceptual model that can be used by SOCs to assess their maturity in each of the six facets and on *Data Driven Security Operations* as a whole. The instantiation should be seen as a quick scan and not be taken conclusively, because there's room for improvement, but despite this we still think it provides a way for better understanding the conceptual model.

Problem Relevance

The market for security technology is currently full of options and sales are driven by (mostly) exorbitant claims. Many of those claims about *Security Analytics* are related to *machine learning*, *anomaly detection*, *artificial intelligence*, *big data* and *predictive security analytics*. Many of these can certainly be applied in a SOC in preventing and detecting security incidents, but there's only limited shared understanding of what it really means to apply analytics and become a *data driven* SOC. Our work is aimed at increasing this understanding.

Design Evaluation

Both the conceptual model and instantiation are evaluated during interviews with experts. Qualitative data has been gathered from each of the interview sessions, which have been summarized in their entirety: all of the questions, remarks and opinions have been recorded. Based on the requirements that had been set in advance, common remarks have been constructed and analyzed to evaluate the results of our work.

Research Contributions

The main contributions of this research are the conceptual model and its instantiation. Jointly they provide a comprehensive

view of *Data Driven Security Operations*, which increases the understanding organizations have of the concept. The instantiation further provides a SOC with concrete advice and recommendations to improve their capabilities, possibly resulting in more effective security operations.

Research Rigor

The conceptual model is based on several types of sources, among which are scientific literature, market reports, vendor marketing material, conference material and survey results. Existing literature addressing (Big Data) Security Analytics (SA) has not been taken as the leading material, as it lacked the practical aspects of security operations. Because of this the research could be labeled as not standing on strong theoretical foundations. That is why we've taken a well-established approach to the construction and evaluation of the artifacts.

Design as a Search Process

The results of this research are very much the result of the search for a common understanding for *Data Driven Security Operations*. Although we provide our own vision and do not claim we have the right answer, we think we have combined many different views into a single, comprehensive model that should help SOCs realizing what it takes to become *data driven*. The model will likely have to change over time, because information security is constantly evolving, so it will remain a search process in the near future.

Communication of Research

The results of the research are primarily aimed at SOCs, and more specifically, the people working in and managing the SOC. This thesis forms the main medium for communicating the results, but the personalized reports for SOCs are better suited to the people in the SOC. The report contains a comprehensive description of the conceptual model along concrete recommendations for the SOC, which convey the message of our thesis in a more concise way.

7.4 DISCUSSION

We have analyzed the transcriptions of the interviews performed with five people from four different organizations to evaluate both the conceptual model as well as the instantiation. The requirements that were constructed before the design phase of both artifacts and which are listed in Tables 5.1 and 5.2 were the main guides and codes for the analysis of the transcriptions.

Based on our analysis we conclude that all of the participants found the conceptual model for *Data Driven Security Operations* as presented to form a complete and coherent picture that was easy to comprehend. According to them it indeed describes what the most important things to consider in *Data Driven Security Operations* are. Participants agree on the genericness of the model, but also indicate that it may be too generic, losing some of its value. Most of the participants are employees of organizations that perform security monitoring and incident response for other organizations as Managed Security Services Providers (MSSPs), which were not the target audience, but in their view the model was still applicable to them. Presenting the model in itself, only showing the six facets, would certainly not have been enough according to the participants. The introduction of the facets through a presentation was absolutely necessary.

Our instantiation received less confirmatory evidence to meeting the requirements. All of the participants had remarks with regard to the way questions were asked in the assessment. Most of the questions that we received feedback on were too broad, not clear enough or discussed aspects that were not known by the participant. We also received feedback on the fact that we did not use quantitative measures, which made the assessment harder to repeat with high confidence, adding to lesser usability in practice. Performing the assessment itself did in fact help the organization to better understand what the model describes, which was one of the goals for creating the instantiation. The instantiation could be improved by quantifying the questions, adding more specific questions and showing all of the submitted answers in the report. Although the validity of the concrete recommendations has not been evaluated in itself, this part of the report was received with positive feedback.

8 CONCLUSION

This research was initiated after observing that Security Operations Centers (SOCs) face numerous difficulties when executing their mission to protect their constituency and recognizing that security solution vendors are only marginally supporting them. We performed an extensive literature study to gain a thorough understanding of SOCs, including their day-to-day practices and the challenges involved. We consulted a diverse set of sources, including marketing material, scientific literature, survey results, webinars and conference materials. This comprehensive study allowed us to answer our first research question: *What challenges do SOCs face nowadays?* From our investigation we conclude that SOCs do indeed face numerous challenges in diverse aspects of their duty. We categorized the challenges that we observed into four main themes, encompassing an increasingly complex IT environment, limited business alignment, ever-evolving adversaries and corresponding attacks, and finally, inadequate resources with respect to people and technology. The challenges need to be addressed appropriately in order for the SOC to successfully execute its mission. We also discovered that vendors are pushing many types of security solutions and are continuously improving their solutions, but relying only on those is not an adequate strategy to the SOC's mission, certainly not when taking the previously listed challenges into account.

In this thesis we present a holistic approach to security operations called *Data Driven Security Operations*, in which people, processes and technology are combined to perform security operations driven by data and analysis thereof. We present the conceptual model of *Data Driven Security Operations* as our first contribution while addressing the second research question: *What should a SOC take into account to address these challenges?* We describe how the model was constructed in a bottom-up approach based mainly on literature research while taking into account the challenges that we identified, but we adopted a top-down approach to describe the model to increase comprehensibility. Our model consists of the following six facets: *Situational Awareness*, *Threat Intelligence*, *Detection Methods*, *Re-*

sponse & Investigation, *SOC Staff* and *SOC Infrastructure*, which all involve data and analysis thereof and together form the conceptual model for *Data Driven Security Operations*. We evaluated our model by performing several semi-structured interviews with experts from the field. All of the participants indicated that they didn't miss particular aspects in the model, from which we conclude that the model gives a complete view of *Data Driven Security Operations*, can be considered valid and is applicable to different types of SOCs. The participants also pointed out that they didn't consider merely presenting the model and its six facets on their own was sufficient to understand what *Data Driven Security Operations* is about. An explanation of each of the facets is necessary to get an adequate level of (shared) understanding of the model, because in their current form, some of them are quite broad or abstract.

Now that we had established the conceptual model, we wanted to provide SOCs with a way to relate to the model, thereby addressing our third research question: *How can SOCs position themselves within Data Driven Security Operations?* We created an instantiation that SOCs can use to position themselves on each of the six facets within our model, giving them an overview of how well they are currently doing. The instantiation consists of a front and a back end that together make up the *Data Driven Security Operations Quick Scan*. The assessments has been implemented using Google Forms and consists of Likert items. One or more members of the SOC can fill in each of the Likert items, categorized into the six facets of the conceptual model, after which a score is calculated for each of the facets. The back end creates a radar chart showing the SOC where it fits within the model for *Data Driven Security Operations* and which facets it can improve based on the goal set by the SOC and the resulting score. The final report consists of three parts, the first of which is a management summary including the resulting radar chart, a full description of the conceptual model for *Data Driven Security Operations* and, finally, some concrete recommendations to consider for improving the SOC based on a goal set by the participant and the resulting score. We evaluated the instantiation by performing semi-structured interviews in the same way as we evaluated the model, but this time we also analyzed the remarks made by the participants. In general, the participants were less positive about the instantiation than about the model, stemming mainly from the way questions were posed in the assessment. According to them, im-

proving the questions, perhaps increasing the number of questions and making them more concrete, would contribute considerably to a more usable instantiation. The participating SOCs did indicate that they liked the format of the report however, including the visualization of the (scored) model as a radar chart. They also appreciated the fact that concrete recommendations are included in the report, but these have not been formally evaluated. We conclude that the current implementation is not sufficient at this moment to conclusively position a SOC within the conceptual model for *Data Driven Security Operations*, but does contain promising parts.

We conclude that the conceptual model for *Data Driven Security Operations* is a good starting point for addressing and discussing the challenges that SOCs face these days and what to consider to improve. Our instantiation can be used to get a rough idea about the current status of the SOC with regard to efficient and effective data analysis and gives SOCs some concrete pointers to consider for improvement. The instantiation also increases the understanding of the conceptual model, but should be improved before more extensive usage.

9 FUTURE WORK

Several avenues for future work were discovered during the evaluation of our work. The first example we would like to bring forward has already been discussed to some extent in Chapter 7 and is aimed at improving our instantiation. As described in the aforementioned section, all of the participants indicated that the assessment could be improved by providing more questions, increasing the specificity of questions, introducing quantitative measures or providing an overview of the answers submitted during the assessment. This direction for future work would consist of the creation of, a possibly different, set of questions that have to be integrated into the existing instantiation and evaluation of the instantiation afterwards, which should include a comparison with the current implementation. Because the extensible approach to the current instantiation, changing the (types) of questions can be performed easily, as long as they stay within the six facets. Adding more parts to the current questionnaire and extending the back end with corresponding processing code is a bit more work, but we still consider this quite manageable.

Currently the calculation of the score depends on the answers submitted by a single person within the Security Operations Center (SOC). Although our implementation could also be used by several staff members of the SOC to more accurately assess the current status, it does not provide an easy way to analyze the combined data gathered in such a way. The instantiation could be extended to allow for multiple users to fill in the assessment for a single SOC and then combining the answers to increase the level of support and accuracy the answers provide. Development of more rigid processes around performing the assessment could also improve the validaty and the efficacy of the instantiation.

We evaluated our work based on performing semi-structured interviews with four different SOCs. Most of these are operating as Managed Security Services Providers (MSSPs), and although we intended the model to be generic and applicable to many differently organized SOCs, the validity of the model and verification of and the usability of the instantiation could be improved

by involving other types of SOCs too, including internally operated ones. Introducing additional expert opinions would also increase the validity of the conceptual model as well as the instantiation by exposing them to a more varied set of people.

So far we have discussed the six facets of our model, how they are interrelated and how they all revolve around data on a relatively high level. Although *SOC Infrastructure* is only a single facet in our model, we have observed that a firm technological foundation is absolutely necessary for the members of the SOC to perform their job effectively. Taking into account the challenges SOCs face nowadays, which we described in section 4.1, and the various aspects of handling, storing and analyzing data in an intelligent and efficient manner, raises the demand for scalable, high performance solutions that are able to integrate reliably. Development, prototyping and implementation of a reference architecture for workloads that SOCs (will) face, might be a viable route for research, especially on a practical level. A common data store capable of ingesting and storing a high volume of security data has been proposed by Marty [68], called the *security data lake* of which the *Apache Metron*[8] project is an example. The security data lake would function as a common repository for security data, including logs, flow data, network traffic and contextual data, examples of which include vulnerability scan reports, configuration information and user identities, that could then be used by various types of consumers, including Security Information and Event Management (SIEM) and forensics solutions, to perform different types of analysis. Using a single repository for ingesting and storing data results in less duplicate data and allows security solutions to perform their function by using shared access to the data. The repository should then be leveraged by security solution vendors so that they can focus on creating innovative products on top of it. Investigation into how to deploy and integrate with the security data repository is a potential direction for future research.

Another potential direction for future research is improving the way information is shared within the information security industry and its customers. Unlike the other suggestions for future work, this is not a direct continued development from our instantiation, but instead surfaced during the interviews with two different participants. Communities where information with regard to threats is begin shared, such as Forum for

8 https://cwiki.apache.org/confluence/display/METRON/Metron+Wiki

Incident Response and Security Teams (FIRST) [9] and other (private) communities already exist, but it can be hard to become a part of these for several reasons. As indicated by one of the participants, vetting may be based on a persons status, reputation and affiliation alone, which is not optimal. Several constructions can be realized that improve this matter, such as vetting based only on contributions to the community, for example, but which are most effective has to be researched. Another participant mentioned that he would like to share information to improve security operations, but that this could possibly affect the MSSP's competitive advantage in a negative way. Research could be conducted to look for ways to increase the willingness to cooperate with other organizations with regard to information sharing, for example by providing only certain pieces of information up front before supplying the receiving party with too many details.

9 https://www.first.org/

A INTERVIEW TRANSCRIPTIONS

This appendix contains the full transcriptions of the interviews performed to validate the conceptual model as well as its instantiation. Each of the following subsections provides a short introduction about the person(s) and the organization they belong to which is followed by the full transcription of the interview. The introductory part about is written in English. We created Dutch transcriptions of the full interview contents, because all of the interviews were performed in Dutch, and this approach thus steers clear from wrongly translating the interviewees views.

DIST

The first part of the interview with DIST was performed with Participant A (*PA*) on October 15th, 2015. *PA* is currently an employee of DIST, a Dutch SME operating in The Netherlands, Belgium and Luxembourg, where he is responsible for a variety of tasks within the company. He leads the consultants and support team, implements firewalls at clients and performs security assessments. As a programmer, he contributes to development of internal systems. His interest for computer security surfaced when he was around 14 years old. He was enrolled in and graduated from his computer science studies at the Saxion University of Applied Sciences. After his formal education he obtained several certificates, including Certified Ethical Hacker (CEH) and several product-specific certifications. Currently he is pursuing an Master of Business Administration (MBA) to learn more about the business side of information security. He hopes this will allow him to create a stronger connection between business and Information Technology (IT) security, which is lacking in numerous companies.

Introduction

DIST is al jarenlang een specialist op het gebied van informatiebeveiliging. Het bedrijf is van oudsher distributeur van diverse (beveiligings)oplossingen, waaronder die van Sophos, DriveLock, Webroot en Ocedo. Er wordt een breed scala aan beveiligingsproducten geleverd aan diverse resellers. Binnen het bedrijf word er tijdens implementaties veel kennis opgedaan op het gebied van diverse firewall oplossingen geproduceerd door onder andere Sophos, Cisco, Juniper en Palo Alto. Daarnaast is er onder andere kennis over diverse Data Loss Prevention (DLP) oplossingen, zoals DriveLock en Sophos DLP en verschillende antivirus en anti-malware oplossingen.

Vanuit de resellers die DIST bedient kwam er steeds vaker vraag naar specialistische kennis op het gebied van implementatie en onderhoud, iets waar DIST door de jaren heen veel van ontwikkeld heeft. DIST levert daarom nu ook (managed) security services en assessments aan (eind)klanten in de vorm van consultancy. Het klantenbestand van DIST beslaat verschillende marktgroepen. De meeste diensten worden uitgevoerd bij middelgrote bedrijven met werknemersaantallen tussen de 300 en 2000 werknemers.

DIST hoopt een brug te kunnen slaan en waarde te kunnen creëren tussen informatiebeveiliging en het bedrijfsleven en de bijbehorende bedrijfsprocessen door de noodzaak van technologische oplossingen te onderstrepen met de specialistische kennis van haar werknemers.

Een van de diensten die DIST levert, is het uitvoeren van beveiligingstesten. Bij een network security assessment wordt het volledige netwerk doorgelicht. Daarvoor wordt gebruik gemaakt van verschillende software, zoals Nessus voor vulnerability assessment en Metasploit voor het uitvoeren van exploits op systemen. Er wordt zo een baseline van de huidige staat van beveiliging vastgesteld. Het rapport van de huidige beveiligingsstatus wordt daarna gerapporteerd aan de desbetreffende organisatie, wat gepaard kan gaan met deskundig advies. De daadwerkelijke verbetering van de security posture van een organisatie wordt, omwille van integriteit, vaak niet uitgevoerd door DIST zelf; daar wordt vaak een derde partij voor ingeschakeld. In sommige gevallen, als het een integere aanpak

niet in de weg staat, kunnen DIST en de organisatie beslissen een vervolgtraject op te starten. De rapportages die DIST levert, bevatten onder andere informatie betreffende welke threats er zijn gedetecteerd, welke zwakheden er vastgesteld zijn en op een hoog niveau welke netwerkcommunicatie er heeft plaatsgevonden.

Een van de bronnen voor analyse is de data die gegenereerd wordt door de firewall die binnen de organisatie draait. Alle sessies die tussen verschillende Internet Protocol (IP)-adressen plaatsvinden, worden daarin gelogd. Als het om een zogenaamde *core firewall* gaat, kunnen ook de sessies die binnen het bedrijfsnetwerk worden opgezet, door DIST gelogd worden. De firewall fungeert dan als het ware als een centrale hub binnen het netwerk, en is in staat om alle communicatie te inspecteren.

De data die de firewall genereert, is veelal beschikbaar in een gestructureerd formaat, waardoor deze vrij makkelijk (in delen) te transporteren en te bevragen is door middel van Structured Query Language (SQL). Het overbrengen van de data van binnen de organisatie naar de door DIST beheerde centrale opslagplaats verloopt via een beveiligde proxy-verbinding. DIST maakt gebruik van SQL om kleinere hoeveelheden data te bevragen en maakt daarnaast gebruik van Logstash om grotere hoeveelheden data te kunnen verwerken en te bevragen. *PA* geeft aan dat de performance van Logstash prima te noemen is, en dat die van SQL bij grote hoeveelheden data erg tegenvalt. SQL wordt met name gebruikt om rapportages te kunnen draaien. Logstash wordt ingezet om een grondiger analyse uit te kunnen voeren op grotere hoeveelheden data.

Bij de analyse van de data wordt er gebruik gemaakt van verschillende correlatieregels. Deze zijn van tevoren gedefinieerd. Zo zijn er regels om te matchen op bijvoorbeeld IP-adressen, hostnamen en applicaties. *PA* bestempelt deze correlatieregels als het laaghangend fruit. Er wordt ook in groeiende mate gebruik gemaakt van Threat Intelligence (TI) feeds, en er zijn contacten met verschillende leveranciers van die data. *PA* noemt dat er in veel gevallen wel standaardmanieren zijn om de data in het interne systeem te plaatsen, bijvoorbeeld door het gebruik van Application Programming Interfaces (APIs) en gestandaardiseerde bestandsformaten. Echter blijkt dat er in sommige

gevallen het bepalen van de context waarin bepaalde gegevens geplaatst moeten worden, soms wel lastig is. Een voorbeeld: van IP-adressen is makkelijk vast te stellen waar het om gaat, maar dat is lastiger te doen voor bepaalde statische tekstregels (*strings*): gaat het dan bijvoorbeeld om bestandsnamen, de inhoud of onderwerpregels van email of de inhoud van postdata gegenereerd door een website? Gestandaardiseerde formaten voor het beschrijven en delen van metadata omtrent deze gegevens, zoals Structured Threat Information eXpression (STIX) en Trusted Automated eXchange of Indicator Information (TAXII), waren *PA* bekend en zouden volgens hem zeker een toegevoegde waarde hebben.

De interne analysesystemen bij DIST worden niet enkel gebruikt voor het leveren van rapportages, die voortvloeiden uit de oorspronkelijke werkzaamheden, maar worden daarnaast ook ingezet ten behoeve van real-time monitoring en detectie. De mate van gebruik ten behoeve van real-time monitoring is de laatste tijd sterk gestegen, nadat bleek dat het inzetten van enkel preventieve beveiligingsmaatregelen, zoals firewall en antivirus, niet toereikend bleek te zijn. Het real-time monitoring aspect levert DIST veel werk op, aangezien er veel data wordt verzameld bij de verschillende aangesloten organisaties.

Om kennis op het gebied van security monitoring binnen het bedrijf te delen, wordt momenteel een *wiki* gebruikt. Daarop worden de technische facetten van de werkzaamheden beschreven. Er zijn continue drie personen werkzaam bij real-time monitoring, en elk van deze personen krijgt dezelfde meldingen binnen. Diegene die op dat moment een melding kan oppakken, doet dat, en geeft daarbij aan dat er aan gewerkt wordt. Zo wordt het werk zo veel mogelijk verdeeld. Het onderzoek van een melding verloopt in drie stappen. Een melding komt binnen bij één van de verantwoordelijke werknemers waarna deze het onderzoek start. Daarbij wordt er een zekere creativiteit van de analist verwacht bij het oplossen van en rapporteren over de melding. Daarnaast vind er periodiek een check plaats van de huidige configuratie van bijvoorbeeld de firewall, die op dat moment ook bijgesteld kan worden. Het melden van een security incident verloopt via een senior security consultant.

Validation

The second part of the interview was performed with *PA* on January 18th, 2016. We started with a quick walk-through for the interview, the first part of which was a short presentation about the conceptual model for *Data Driven Security Operations* and each of facets it consists of. Each of the facets was explained while keeping room for discussion, the gist of which has been transcribed in the next section. During the presentation we also looked at the graphical depiction of the conceptual model and *PA* was instructed on how the evaluation would continue, including how scores had to be filled in. Notes have been taken during the time *PA* was filling in the questions, including the short discussions that followed, and we also transcribed these in the following sections. After *PA* completed the test, we discussed the validity and usability in practice of the conceptual model and its instantiation based on several pre-defined questions, which we also transcribed.

The following sections contain the contents of the discussions before, during and after the test, including the presentation of the conceptual model and final discussion.

Presentatie conceptueel model

De eerste stap van de evaluatie is het kort introduceren van het ontwikkelde conceptuele model inclusief een toelichting van de facetten waaruit het model bestaat. Er werd bij elk van de facetten stil gestaan, waarbij er ruimte was voor vragen en opmerkingen van *PA*'s kant. De beschrijving van deze facetten zal uiteindelijk ook beschikbaar gesteld worden bij het eindrapport van het doorlopen van de test.

Het eerste facet waarbij *PA* een vraag had was Threat Intelligence, waarbij het ging om het gebruik van Threat Intelligence dat niet alleen plaatsvindt op operationeel level, maar ook op strategisch en tactisch niveau. *PA* kon zich vinden in de uitleg dat Threat Intelligence op hoger niveau een beter beeld kan geven waar een organisatie op dat moment mogelijk mee te maken kan krijgen en dat er verschil bestaat in de vorm en kwali-

teit van verschillende Threat Intelligence feeds op operationeel niveau.

Alhoewel *PA* het zich kon voorstellen waarom *SOC Staffing* een onderdeel van het model is, vroeg hij hier wel expliciet naar, omdat hij benieuwd was naar mijn redenatie. Hij kon zich vinden in de uitleg dat de werknemers binnen het SOC niet alleen het huidige werk moeten aankunnen, maar zich ook continue moeten blijven ontwikkelen. Deze claim wordt onder andere gesteund door verschillende SOC maturity modellen. Daarnaast speelt het feit dat de security operators niet alleen technische kennis moeten hebben van monitoring een rol: security operators moeten steeds vaker meer weten van andere takken van sport, zoals statistiek, Threat Intelligence en

Bij het facet *SOC Infrastructure* legde ik uit dat het hier ging om hoe goed de technologie de huidige SOC functies ondersteunt. Het gaat daar om uiteenlopende zaken, zoals performance, schaalbaarheid en de mogelijkheid tot uitbreiden en compatibiliteit van de infrastructuur. *PA* kaart het punt van de beveiliging van de SOC infrastructuur aan, waarop ik antwoordde dat dit inderdaad een rol speelt in mijn model en terugkomt in de test, met daarbij een korte beschrijving van dit aspect van de SOC Infrastructuur.

Testdoorloop

Bij het invullen van de eerste pagina van de test, het aangeven van de weging, hebben we het kort gehad over waarom *PA* een bepaalde weging toekende aan de verschillende facetten. *PA* gaf aan een lichte voorkeur (4) in weging te hebben met betrekking tot *Response & Investigation*, wat volgde uit het feit dat DIST op dat gebied volgens hem het meeste te ontwikkelen heeft.

De overige weegfactoren kende *PA* een 3 op een schaal van 1 tot en met 5 toe, wat inhoudt dat deze een voordeel noch een nadeel ontvangen in de uiteindelijke berekening van de scores.

Na het toekennen van weegfactoren aan de verschillende facetten werden scores op een schaal van 1 tot 5 toegekend aan de huidige staat. Tijdens het vaststellen van deze waarden hebben

we gesproken over waarom *PA* ergens een bepaalde waarde toekende. Bij het facet *Situational Awareness* gaf *PA* aan dat DIST daar zeker mee bezig is, daar ook de noodzaak van inzien, maar dit nog niet volledig uitgedacht en -ontwikkeld hebben; de toegekende waarde is 2. Omdat DIST als Managed Security Services Provider (MSSP) opereert komt het namelijk vaak voor dat ze bij klanten geen toestemming krijgen om alle data te verzamelen die nodig is om een compleet beeld van de IT omgeving van die klant te kunnen krijgen.

DIST maakt reeds gebruik van Threat Intelligence feeds en houdt ook het nieuws in de gaten om vast te stellen wat er speelt om zo hun klanten proactief te kunnen beschermen tegen nieuwe ontwikkelingen. Zodra er nieuws over nieuwe exploits beschikbaar is ontwikkelt zij daar herkenningsregels voor die vervolgens voor klanten worden geactiveerd.

Ook het facet *Detection Methods* scoort een 3. Naast het gebruikmaken van de regelsets die geleverd worden door de producenten van de producten die DIST bij haar klanten plaatst, ontwikkelt men binnen het bedrijf namelijk ook aan eigen regelsets. Daar zitten bijvoorbeeld regels bij voor het detecteren van afwijkend gedrag bij authenticatie van gebruikers.

Het facet *Response & Investigation* wordt door *PA* al gewaardeerd op een 4, maar is nog wel aan verbetering toe. De processen zijn nog voornamelijk ad-hoc en intern niet gestructureerd genoeg. Bij de communicatie richting klanten bestaan er wel gestructureerde processen. Wat betreft proactief onderzoek zoekt DIST op basis van Threat Intelligence feeds en nieuws naar mogelijke incidenten bij haar klanten. Dit was bijvoorbeeld het geval voor de backdoor in Juniper ScreenOS die in december 2015 publiekelijk aan het licht kwam.

SOC Staffing werd een 3 toegekend, mede omdat er mensen aan het werk gezet worden die een vooropleiding (technische) informatica hebben gevolgd. Daarnaast biedt DIST de mogelijkheid om product-specifieke certificaties na te streven en om certificaten zoals CEH en Certified Information Systems Security Professional (CISSP) te behalen. Binnen DIST zelf is er brede kennis aanwezig van information security, omdat dit is waar het bedrijf zich in heeft gespecialiseerd, en niet zozeer in security mo-

nitoring an sich, met bijbehorende voordelen met betrekking tot het bedienen van klanten.

PA geeft aan dat de huidige *SOC Infrastructuur* goed voor elkaar is en waardeert dit facet op een 4. De benodigde data kan momenteel snel en volledig genoeg binnengehaald worden, en dit gebeurt ook op veilige wijze, door het dataverkeer af te schermen via verschillende proxies. Er zijn audit trails aanwezig die beschrijven wie er bepaalde data heeft ingezien. Om de schaalbaarheid van opslag en analyse te garanderen, wordt er gebruik gemaakt van Elasticsearch, Logstash en Kibana (ELK-stack).

Aandachtspunten tijdens testdoorloop

- Bij het starten van de vragenlijst vraagt *PA* zich af hoe hij de vragen zou moeten beantwoorden. Moet dat vanuit het perspectief van een MKB bedrijf, wat DIST en haar klanten zijn, of vanuit een veel grotere onderneming.
- *PA* geeft aan dat de manier van vraagstelling eigenlijk eist dat verschillende personen binnen hetzelfde bedrijf, maar op verschillend niveau, dezelfde vragen zouden moeten beantwoorden, of verschillende delen daarvan.
- Het concept van een *threat model* zou duidelijker uitgelegd mogen worden.
- Het verkrijgen van een high-level view zou misschien alleen van toepassing moeten zijn op de meest kritische bedrijfsprocessen.
- Strategisch, tactisch en operationeel level zijn mogelijk niet voor iedereen duidelijk en eenduidig.
- Sommige vragen, zoals het maken en delen van Threat Intelligence bevatten twee elementen die eigenlijk los van elkaar gezien moeten worden. DIST doet bijvoorbeeld wel zelf aan het maken van nieuwe IOCs, maar deelt deze niet openbaar of privaat met andere bedrijven.
- LDA wordt niet gedaan, focus op specifieke entities, nog in ontwikkeling
- Bij procedures voor reactief en proactief noemt *PA* dat die er gedeeltelijk wel zijn, maar dat deze niet volledig uitgewerkt zijn tot procedures. De desbetreffende vragen zouden dus los van elkaar gezien kunnen worden.

- In het MKB zouden er over het algemeen minder formele procedures bestaan voor bepaalde activiteiten. Vragen die hier dieper op ingaan zouden gecombineerd kunnen worden in een enkele.

Discussie

Na afloop van het doorlopen van de test is er gediscussieerd over het conceptueel model en de praktische implementatie daarvan, inclusief ingevuld model. Er is in eerste instantie niet verder gepraat over de inhoud van het gegenereerde rapport dat voorstellen ter verbetering bevat.

Het conceptuele model, bestaande uit zes facetten vond *PA* een compleet en coherent plaatje vormen. Hij gaf aan dat hij zo snel geen omissies in het model had opgemerkt, en dat de verschillende facetten op voldoende hoog level waren beschreven dat deze toepasbaar zouden zijn in de praktijk en op verschillende niveaus. *PA* stipt aan dat het risico van dit model is dat het misschien wel te breed is, maar dat er anderzijds ook weer iets voor te zeggen is als een bepaald facet niet benoemd wordt. Dit is met name het geval voor het facet SOC Staffing: waar de andere facetten voornamelijk om techniek draaien, valt deze op dat gebied een beetje buiten de boot. Het toevoegen ervan zorgt echter wel voor een completer plaatje van waar organisaties rekening mee moeten houden.

In principe was het voor *PA* duidelijk waar het bij de zes facetten om ging, al voegde de introductie van het model wel kennis toe op het facet *Situtational Awareness* en *Threat Intelligence*. De definitie van TI in het conceptuele model bleek breder te zijn dan *PA* zelf in gedachten had. Daarnaast behoefde het facet SA wel een kleine uitleg, maar begreep *PA* waar het om ging. Er werden geen zwakke punten vastgesteld met betrekking tot de facetten an sich. Uiteraard zijn er wel zaken die genuanceerd kunnen worden, waarbij misschien wel het belangrijkste het verschil tussen MKB en grotere onderneming van belang is.

De grafische weergave van het model na afloop van de testdoorloop vond *PA* een goede weergave. Het is meteen duidelijk waar het om gaat, alhoewel de *perceived level* misschien wel aangeduid moet worden met een (kleine) legenda. Naast de grafische weergave stelt *PA* voor dat een tabel met de ingevulde

waarden ook handig kan zijn, om zo de test beter herhaalbaar te maken: de ingevulde waarden zijn dan immers nog bekend. Op de vraag of de (optionele) notities aan het eind van elk facet daar ook aan zouden helpen, antwoordt hij bevestigend. De mogelijkheid om dit per vraag te kunnen doen is zijns inziens wellicht minder praktisch, want dan zou het in totaal erg veel kunnen worden.

De bruikbaarheid, herhaalbaarheid en vergelijkbaarheid van de huidige implementatie zou in de praktijk veel groter worden als de verschillende facetten beter meetbaar zouden worden. Enerzijds kunnen specifiekere vragen daar een antwoord op zijn: zodra een vraag specifieker gesteld wordt, is er minder ruimte voor interpretatie door diegene die de test afneemt. Sommige vragen zijn momenteel zo breed of bevatten meerdere delen, zodat het lastig is om ze heel precies in te vullen. Door die vragen op te delen ontstaan er weliswaar wel meer, en zou iemand langer doen over de test, maar wordt het wel concreter waar het bij zo'n vraag om gaat. Er zou volgens *PA* ook gekeken kunnen worden naar vragen met een meer kwantitatieve insteek, bijvoorbeeld het noemen van aantallen of percentages. Door dit soort vragen toe te voegen zouden er minder mogelijkheden zijn om de antwoorden te beïnvloeden, omdat er beter te controleren is wat er is ingevuld.

De huidige opzet van de praktische implementatie zou beter tot zijn recht komen in grotere organisaties om de huidige situatie weer te geven dan in het MKB. Het model zelf, bestaande uit de facetten en de structuur, zet echter al wel aan tot denken, waar de vragen een hulpmiddel bij zijn. *PA* heeft de rapportage in grote lijnen bekeken, en die leek hem in ieder geval in deze vorm al praktisch bruikbaar. Het type model is volgens hem ook zeer bruikbaar om een vergelijking tussen organisaties en/of de eigen organisatie op een ander moment uit te kunnen voeren.

De insteek van het model, het bepalen van de *maturity* van organisaties op verschillende facetten, is volgens *PA* ook te relateren aan het INK model en bijbehorende methoden. Dit gaat dan met name om het breder uitzetten van dezelfde vragen over verschillende lagen in een organisatie, om zo een accurater beeld te kunnen krijgen van de verschillende facetten op verschillende niveaus. Bij het INK model is het gebruikelijk om

20% van de mensen binnen een organisatie te laten deelnemen en daar conclusies op te baseren, om zo niet te afhankelijk te worden van een enkele persoon binnen die organisatie.

MSSP1

At MSSP1 the first interview was performed with Participant B (*PB*), the CEO of the company on January 6th, 2016. *PB* has many years of experience in the area of information security. As one of the directors of CONS he is responsible for the how the business operates and a consultant in information security. Within CONS he has gained a lot of experience in designing and developing security monitoring capabilities and the creation of Security Operations Centers (SOCs). He also worked at the SOC operated by Ziggo, a Dutch telecommunications provider, and was active within the Nationaal Cyber Security Centrum (NCSC). In addition to his responsibilities at CONS, *PB* is also in charge of MSSP1. He has a lot of knowledge about and experience in Security Information and Event Management (SIEM), Sherwood Applied Business Security Architecture (SABSA), en various standards related to information security, including ISO27001 and NEN7510, and he is a Certified Information Systems Security Professional (CISSP).

Introduction

MSSP1 is een bedrijf gevestigd in het oosten van Nederland dat zich richt op het leveren van security monitoring diensten. Het bedrijf is daar halverwege 2015 mee gestart, na het zien van een gat in de markt. Binnen CONS, waarvan MSSP1 een onderdeel is, bestond namelijk het beeld dat er vraag was naar security monitoring diensten binnen middel tot grote organisaties, zoals gemeenten, onderwijs- en zorginstellingen. Er bestonden echter niet veel organisaties die dit in eigen beheer konden realiseren of voor andere bedrijven uit handen konden nemen. Vanuit CONS was er al veel kennis aanwezig op het gebied van compliance monitoring en security testing, bijvoorbeeld in de vorm van *penetration tests* en hadden enkele werknemers uitgebreide ervaring opgedaan met het opzetten van SOCs binnen zeer grote omgevingen, zoals KPN, ING en Thales. Het leveren van security monitoring is typisch een dienst die door Mana-

ged Security Services Providers (MSSPs) geleverd wordt, onder welke noemer MSSP1 zich dus ook onder kan scharen.

Op dit moment bedient MSSP1 een gering aantal klanten en is zij bezig om haar diensten bij een vrij grote organisatie uit te rollen. Het klantenbestand bestaat nu voornamelijk uit educatieve instellingen.

MSSP1 maakt gebruik van LogPoint om haar security monitoring diensten aan te kunnen bieden. LogPoint is een SIEM en log management oplossing van een relatief kleine speler op SIEM gebied van Deense bodem. Er is om meerdere redenen gekozen voor LogPoint, waarvan de mogelijkheid voor multi-tenancy en het feit dat LogPoint een relatief kleine speler is, de belangrijkste zijn. Multi-tenancy, oftewel, het kunnen aanbieden van diensten aan verschillende klanten vanuit een enkele installatie, zorgt ervoor dat MSSP1 meerdere klanten kan bedienen zonder daarvoor een geheel nieuwe instantie van de harden/of software neer te zetten. Het feit dat LogPoint een relatief kleine, en onafhankelijke (van HP, IBM, en soortgelijke conglomeraten) partij is, zorgt ervoor dat er snel gehandeld kan worden als MSSP1 tegen eventuele problemen aanloopt. Daarnaast is de licentiestructuur, die gebaseerd is op het aantal machines dat gegevens stuurt, erg overzichtelijk en zorgt dat voor een duidelijk kostenplaatje voor de klant. *PB* voegt daar aan toe dat in gevallen waar andere SIEM oplossingen gebruikt worden, bepaalde gegevens niet opgeslagen worden om de kosten te drukken.

Bij de klanten die MSSP1 bedient, plaatst zij een installatie van LogPoint die de gegevens voor die specifieke organisatie aggregeert, de *collector* genaamd. Tijdens de implementatie wordt er gekeken welke gegevens er verzameld moeten worden, en wordt de installatie naar specifieke eisen en wensen ingericht. Die installatie vindt plaats op een fysieke machine, aangezien er problemen ondervonden werden bij het draaien van LogPoint in een virtuele machine. Op de lokale instanties van LogPoint wordt het grootste deel van de log gegevens voor langere tijd opgeslagen en geanalyseerd. De machines zijn dan ook uitgerust met voldoende opslagruimte om een vooraf ingestelde hoeveelheid tijd aan gegevens op te kunnen slaan. De resultaten van die eerste analyses, gebaseerd op vooraf ingestelde regelsets, worden doorgestuurd naar de centrale installatie die

zich in een datacentrum bevindt en vanaf de bedrijfslocatie toegankelijk is. MSSP1 krijgt zo een centraal beeld van alle organisaties die aangesloten zijn op LogPoint en kan de gehele dataset gebruiken voor analyses. De collector instanties zijn door MSSP1 op afstand te beheren, waarvoor gebruik gemaakt wordt van twee verschillende Virtual Private Network (VPN) verbindingen. Een van de verbindingen is voor het beheer van de fysieke machine en de andere verbinding wordt gebruikt voor het beheer van de LogPoint installatie om zo bijvoorbeeld nieuwe regels aan te kunnen maken.

Het gebruik van LogPoint biedt MSSP1 veel voordelen ten opzichte van andere SIEM oplossingen, zoals ArcSight (HP) en QRadar (IBM), maar mist nog wel enkele functionaliteiten die het bedrijf graag had willen zien. Zo is LogPoint, net als andere SIEM oplossingen, afhankelijk van een regelset om ongewenste en verdachte gebeurtenissen te kunnen detecteren. Het onderhouden van zo'n regelset is veel werk, dat bovendien gevoelig is voor fouten of onvolledigheden. Daarnaast biedt LogPoint wel enkele mogelijkheden om de gehele dataset te kunnen doorzoeken, maar zijn die mogelijkheden wel beperkt tot wat de gebruikersinterface biedt. Om dit voor een deel op te lossen maakt MSSP1 daarom gebruik van de ELK-stack (Elasticsearch, Logstash en Kibana) om de data makkelijker te kunnen bevragen, maar moet daarvoor wel eerst de gegevens overzetten van LogPoint naar ELK. *PB* geeft aan dat er op dit moment echter nog niet genoeg gedaan wordt naar het zoeken in de data om verdachte zaken die niet door LogPoint gevonden worden te kunnen vinden, en als dat al gebeurt, dit geen vaste, herbruikbare methode(n) betreft.

Validation

The second part of the interview at MSSP1 was performed on January 21st, 2016. We started with a quick walk-through for the interview. The conceptual model was presented to *PB*, together with each of the facets it's composed of. Each of the facets was discussed to get a shared understanding of them. During the presentation we also had a look at the graphical depiction of the conceptual model. During the test notes have been taken to capture *PB*'s thoughts about the model and the test. These have been transcribed in the following sections. After the test we had a discussion about the validity and practi-

cality of the conceptual model and the current implementation based on pre-defined questions. These have also been transcribed below.

The next part consists of the transcriptions of the discussions before, during and after the test, including the presentation and final discussion about the model and implementation.

Presentatie conceptueel model

Als eerste kreeg *PB* de verschillende facetten van het conceptuele model te zien waarbij een korte uitleg werd gegeven van wat de facetten inhouden. De presentatie van het conceptuele model verliep eigenlijk geheel volgens schema, en *PB* gaf aan de verschillende facetten te begrijpen.

Testdoorloop

De test is in vergelijking met de vorige afname op een enkel groot punt grotendeels gelijk gebleven. Aan het begin wordt nog steeds de belangrijkheid (weging) van elk van de facetten bepaald, maar de inschatting van het huidige niveau is verschoven naar het einde van de test. Zie ... waarom dit aangepast is. Daarnaast wordt een doelstelling bepaalt aan het einde van elk van de facetten. Afgesproken werd om in het geval van MSSP1, daar zij een MSSP is, de antwoorden op basis van het gehele klantenbestand gegeven zouden worden.

Bij het invullen van de belangrijkheid van elk van de facetten voor MSSP1, hebben *PB* en ik kort besproken waarop de keuze voor een bepaalde waarde gebaseerd was. *Situational Awareness* en *Threat Intelligence* werden beiden gewaardeerd op een 4, wat voortkomt uit het feit dat MSSP1 graag een goed inzicht heeft in de huidige status van de IT omgeving en wat daarbuiten gebeurt. Aangezien het detecteren van aanvallen en inbraken de kern van de werkzaamheden van MSSP1 is, wordt *Detection Methods* op 5 gezet.

MSSP1 voert niet zelf heel diepgaande onderzoeken uit op het moment dat een incident gedetecteerd wordt. Dit wordt opgevangen door een andere partij, wat CONS, het moederbedrijf van MSSP1, zou kunnen zijn, of een geheel andere partij. Dit is waarom de het facet *Response & Investigation* gewaardeerd

wordt op een 3 wat betreft belangrijkheid. De overige facetten, *SOC Staffing* en *SOC Infrastructure* kregen door *PB* een 5 toegekend, omdat deze voor MSSP1 naast *Detection Methods* het belangrijkste zijn om haar diensten aan klanten te kunnen leveren. *PB* noemt daarbij nog dat het werven van ervaren mensen momenteel een van de grootste moeilijkheden vormt voor MSSP1.

Aandachtspunten tijdens testdoorloop

- Op het gebied van rapportages m.b.t. de huidige beveiligingsstatus kan MSSP1 nog veel verbetering realiseren.
- Als MSSP is het lastig om te bepalen of je voldoende inzicht hebt in de omgeving van de klant. De klant heeft immers controle over de IT omgeving, en daarmee een overzicht van de primaire systemen.
- MSSP1 denkt wel na over mogelijke scenario's, maar doet dat niet op structurele wijze. Daarnaast heeft zij niet voldoende zicht op de IT omgeving om simulaties uit te kunnen voeren.
- Inventarissen van vulnerabilities, patches en configuratiemanagement systemen zijn over het algemeen nooit volledig op orde.
- Bij het onboarden van een nieuwe klant wordt een network topologie geleverd door de klant, waarop MSSP1 moet vertrouwen dat deze up-to-date en accuraat is. Ook verloopt het proces van het aannemen van een nieuwe klant in stappen, waarbij steeds één of meerdere onderde(e)l(en) van de bedrijfsprocessen worden ingericht voor monitoring. Configuratie en log collectie vindt dus gefaseerd plaats.
- Het kunnen bekijken van de huidige staat van het gehele netwerk is vaak niet mogelijk. Dit heeft o.a. te maken met het voorgaande punt, namelijk dat niet alle processen tegelijkertijd aangesloten hoeven te zijn. Naast het inrichten van monitoring op de primaire bedrijfsprocessen, worden wel altijd ook andere sensoren aangesloten, waaronder bijvoorbeeld firewalls en Intrusion Detection Systems (IDSs).
- *Contextualization* was *PB* in eerste instantie niet helemaal duidelijk. het toevoegen van bedrijfsspecifieke informatie

aan de data collectie, zoals gebruikersrollen, past MSSP1 niet uitgebreid toe.

- De vraag over het niveau van requirements, evaluatie, etc. van Threat Intelligence is vrij breed, waardoor deze lastig echt te duiden is op een specifieke waarde.

- MSSP1 is in gesprek met Quarantainenet over het delen van Threat Intelligence, maar nog niet echt daarbuiten. Bij CONS is daar aanmerkelijk meer kennis over aanwezig. Daarnaast wordt er gekeken naar een aansluiting bij Forum for Incident Response and Security Teams (FIRST).

- Er zijn nog veel handmatige processen rondom Threat Intelligence.

- Honeypots zijn in een eerder stadium wel toegepast, maar op dit moment niet meer. Verder zijn de aspecten rondom *Detection Methods* volgens *PB* goed ingericht.

- *Response & Investigation* is een facet waar op dit moment bij MSSP1 niet de focus op ligt. Het bedrijf richt zich primair op het monitoren van klanten, en schakelt andere partijen, zoals CONS in, zodra er echt ingegrepen moet worden. Over de gehele lijn kregen de vragen dan ook een lage waarde toegekend.

- Processes rondom educatie, certificering en samenwerking zijn goed ingericht bij MSSP1. Gestandaardidseerde workflows kunnen echter beter, maar zijn in ontwikkeling.

- Over de gehele linie heeft meer dan voldoende MSSP1 kennis in huis, maar op forensisch gebied is dat wat minder. Gezien de prioriteit van MSSP1 voor monitoring, is dat niet direct een heel groot probleem. *PB* voegt er wel aan toe dat meer kennis altijd welkom is.

- Over het algemeen zit de monitoring infrastructuur van MSSP1 goed in elkaar.

- Linked data analyse is mogelijk, maar wordt op dit moment nog niet toegepast.

- De veiligheid van de eigen infrastructuur is nog niet formeel getest, maar de eigen apparatuur wordt wel gemonitord.

- *PB* mist specifiek een vraag over de continuïteit en beschikbaarheid van de monitoring infrastructuur.

Discussie

Aan de hand van een aantal vooraf vastgestelde vragen werd er gediscussieerd over het conceptueel model en de huidige implementatie, inclusief de berekende waarden voor MSSP1. In eerste instantie zijn de concrete voorstellen ter verbetering niet stuk voor stuk doorgenomen; daar zou een nieuw rapport voor gegenereerd worden.

PB vond dat de zes verschillende facetten een goed beeld gaven van waar het momenteel om draait in security monitoring. De zes facetten geven volgens hem een compleet en coherent beeld. Bepaalde facetten zijn voor MSSP1, als MSSP, uiteraard belangrijker dan andere, waardoor tijdens het doorlopen van de test sommige vragen veel minder van toepassing waren op de werkzaamheden van MSSP1 dan andere vragen. De meerwaarde als MSSP voor organisaties komt volgens *PB* voort uit het feit dat de security operators in het SOC meer gedegen zijn in hun vakgebied: zij zijn immers constant bezig met het monitoren van netwerken, en het is niet slechts een onderdeel van de taken.

Wat volgens *PB* concreter aan het licht mag komen in het model is de fysieke beveiliging en monitoring daarop. Voor sommige organisaties is dit inderdaad van belang, omdat het threat model daar ook fysieke toegang als concrete aanvalsvector wordt beschouwd. Het geldt ook voor MSSP1 zelf, aangezien zij toegang heeft tot mogelijke vertrouwelijke data, en toegang tot het bedrijfsgebouw niet direct mag betekenen dat men daar ook bij komt.

De visualisatie van het model was voor *PB* een herkenbaar beeld, en het was volgens hem direct duidelijk waar het om ging. De verschillende facetten van het conceptuele model worden er duidelijk in getoond, en het oogt intuïtief. Het geeft volgens hem ook goed weer waar er verschillen zijn met de

doelstelling, alhoewel die op dit moment misschien allemaal erg hoog ingezet werden.

PB geeft aan dat de huidige vorm van de visualisatie zeker bruikbaar is om te vergelijken, maar dat bij een herhaling van de test mogelijk wel andere antwoorden gegeven worden. Het voorstel om verschillende, nu nog erg brede, vragen op te delen zou een oplossing kunnen zijn, maar zou volgens *PB* wel zorgen voor een veel langere tijd om de test door te lopen, en dat dit mogelijk leidt tot afdwalen. Daar moet dus een gulden middenweg in gerealiseerd worden. Om de test te kunnen reproduceren zou een overzicht van alle gegeven opties in bijvoorbeeld een tabel een waardevulling zijn in het huidige rapport. Het gebruik van beter kwantificeerbare vragen zou volgens *PB* niet heel veel verschil uit maken, tenzij de benodigde gegevens om dat soort vragen te beantwoorden echt aanwezig is. Met *PB* werd nog besproken om meerdere mensen binnen dezelfde organisatie te vragen de test in te vullen, en dan een gemiddelde te nemen van alle antwoorden, of om verschillende onderdelen aan verschillende verantwoordelijke te vragen. Hij kon zich vinden in deze manier van beter meetbaar maken van de test.

Ondanks de haken en ogen aan de huidige vraagstelling en herhaalbaarheid van de test, vindt *PB* het model en de rapportage in deze vorm zeker bruikbaar in de praktijk.

CERT

The entire evaluation with CERT was performed on February 26. The interview was performed with Participant C (*PC*), who is in charge of the Computer Emergency Response Team (CERT). *PC* started working at the educational institution as a network administrator about 25 years ago. Back then, he already had the idea that security was an important aspect of IT, and he has been steadily developing himself towards security oriented tasks. Today he is the security manager for the educational institution, equivalent to 0.5 FTE. Before he started working at his current employer, *PC* worked at the HTS Arnhem (now HAN) and for several years at Philips.

Introduction

CERT heeft tot doel om de informatiebeveiliging op onderwijsinstelling te bewaken. Als onderdeel van de IT-dienstverlener van de onderwijsinstelling, monitort het CERT het netwerk en grijpt het in bij beveiligingsincidenten. Die incidenten kunnen gemeld worden door werknemers of studenten, maar kunnen ook voortkomen uit de monitoring van het netwerk. Het werkveld van CERT beslaat alle services, software en hardware die worden aangeboden via de IT-dienstverlener van de instelling. Twee specifieke werkzaamheden die het CERT uitvoert, zijn het blokkeren van apparaten op het netwerk en het van de perso(o)n(en) die verantwoordelijk zijn voor misbruik.

CERT werkt samen met andere partijen om haar werkzaamheden te verrichten. Zo werkt zij onder andere samen met SURFnet, die als Internet Service Provider (ISP) voor de externe netwerktoegang verantwoordelijk is en verschillende diensten aanbiedt om onderzoekers en studenten efficiënt samen te laten werken. Via SURFnet wordt er ook informatie uitgewisseld die betrekking heeft op informatiebeveiliging. Ook met andere onderwijsinstellingen en de bijbehorende CERTs wordt informatie uitgewisseld op het gebied van informatiebeveiliging. Zodra er een incident plaatsvindt waarvoor CERT niet direct verantwoordelijk is, zet zij deze door naar de verantwoordelijke beheerder.

CERT heeft een vrij goed beeld van de IT omgeving die zij beschermt. Het gaat daarbij om de hardware en software en bijbehorende configuraties die door de IT-dienstverlener geleverd wordt; systemen die door medewerkers zelf geplaatst worden, zijn dus niet niet of slecht inzichtelijk voor CERT. Ze reageert wel op meldingen die voortkomen vanuit niet-standaard systemen. Door het een centraal ingericht systeem, zijn de identiteiten van gebruikers bekend. Met behulp van andere tools kan CERT deze data gebruiken om in te zien wie er op welk moment toegang heeft gehad via Eduroam of Virtual Private Network (VPN). Ook maakt CERT gebruik van actieve netwerk scans om het gehele netwerk in kaart te brengen, inclusief relaties onderling, om zo een compleet beeld van het netwerk te kunnen realiseren.

Een deel van de meldingen komt binnen via automatische systemen. Hiervoor wordt onder andere gebruik gemaakt van Qua-

rantainenet, die oplossingen biedt voor het automatisch detecteren en isoleren van gecompromitteerde machines. Ook wordt er gebruik gemaakt van antivirus software op de systemen die door de IT-dienstverlenerbeheerd worden. In combinatie met het feit dat er een behoorlijk accuraat beeld van de huidige IT omgeving gerealiseerd kan worden, zorgt dit voor een tijdige aanpak van incidenten.

De werknemers die actief zijn binnen CERT vallen onder de IT-dienstverlener, en vervullen hun werkzaamheden die betrekking hebben op security dan ook op part time basis. Er wordt in brede zin aandacht besteed aan de ontwikkeling van werknemers, niet alleen specifiek op het gebied van security, maar in een breder geheel. De werknemers hebben op hun eigen deelgebied (bijvoorbeeld Linux, Windows, netwerktechnologie) zeer specifieke kennis, ook op het gebied van security. Kennisdeling tussen de verschillende mensen verloopt tegenwoordig nog wel ietwat stroef.

Validation

As described before, the validation was also part of the interview on February 26, 2016. After a short walk-through for this part, the conceptual model for *Data Driven Security Operations* was presented to *PC*. Each of the six facets was explained with room for discussion. After that, the graphical version of the model was presented, including how a score would be computed on the model. Notes have been taken during the test to capture *PC*'s thoughts about the questions which are described in the next subsections. At the end of the test we had a discussion about the validity and practicality of the conceptual model and the current implementation based on pre-defined questions. These have also been summarized below.

The next part consists of the transcriptions of the discussions before, during and after the test, including the presentation and final discussion about the model and implementation.

Presentatie conceptueel model

Het eerste onderdeel was de presentatie van het model voor *Data Driven Security Operations*. Elk van de zes facetten werd

kort mondeling toegelicht. Bij dit onderdeel had *PC* geen vragen en/of opmerkingen.

Testdoorloop

Voordat met de doorloop van de test gestart werd, is uitgelegd hoe de test in elkaar steekt. Het uitgangspunt voor het beantwoorden van de vragen is het CERT, dat als onderdeel van de IT-dienstverlener functioneert.

Bij het invullen van de doelstelling van elk van de facetten gaf *PC* aan dat het College van Bestuur een standaard procedure voor dit soort assessments heeft vastgesteld, waardoor elk van de doelstellingen op 3 werd bepaald. Na de eerste doorloop worden de verschillen vastgesteld met de doelstelling en wordt er gekeken naar waar er verbeterd kan worden. Dan kunnen ook de doelstellingen aangepast worden voor een eventuele volgende doorloop van een assessment.

Aandachtspunten tijdens testdoorloop

- CERT is momenteel bezig om (kritieke) bedrijfsprocessen in beeld te brengen. Van daaruit zal ook het risicoprofiel vastgesteld kunnen worden.
- De huidige rapportages geven wel inzicht, maar zijn zeker niet volledig.
- Het belang van Threat Intelligence wordt begrepen, maar wordt nog niet structureel toegepast. Dit gebeurt voornamelijk binnen lopende projecten. Het wordt op projectniveau toegepast, maar niet op organisatorisch niveau.
- Heuristiek en anomalie detectie worden toegepast, maar slechts op een deel van al het netwerkverkeer.
- Detectiemethoden worden regelmatig, op ad-hoc basis gereviewed.
- Continuïteit in incident response is niet eenduidig. Bij CERT is er wel 24x7 een team aanwezig om incidenten af te handelen, maar het is niet hun enige taak.
- De wijze waarop de vraag over de regelmatigheid van reactief onderzoek naar een incident wordt gesteld, is *PC* niet direct duidelijk.

- Er wordt op ad-hoc basis proactief onderzoek uitgevoerd op het netwerk. Een voorbeeld is het scannen van het netwerk op de aanwezigheid van Raspberry Pi's met een standaard wachtwoord nadat een gecompromitteerde Raspberry Pi gelokaliseerd is op het netwerk om dit in de toekomst te kunnen voorkomen.

- Er is ruimte voor de ontwikkeling van kennis en kunde. Dit is niet specifiek op security gericht, maar verdeeld over het hele scala van IT diensten.

- Er zijn middelen om kennis te delen, maar daar wordt niet altijd volledig gebruik van gemaakt.

- De vraag over Open Source Intelligence en Reverse Engineering bestaat volgens *PC* uit twee totaal verschillende delen. Er is volgens hem wel wat voor te zeggen, maar het maakt de beantwoording van de vraag wel lastiger.

- De vraag over de kwaliteit van de metrics voor security operations roept ook vragen op. Allereerst wilde *PC* weten waar deze metrics dan om draaien. Uiteindelijk wordt deze vraag beantwoordt met een lage score, omdat er op dit moment eigenlijk geen sprake is van aanwezigheid van deze metrics.

- Aan het inrichten van de beveiliging van de Security Operations Center (SOC) infrastructuur wordt momenteel nog gewerkt. Er komt een aparte server voor security operations. Bepaalde onderdelen zijn gebouwd door andere onderdelen van de IT-dienstverlener, waar *PC* zo af en toe nog wel zijn vragen bij heeft. Ook is bepaalde data structureel toegankelijk voor personen die daar slechts incidenteel toegang toe nodig hebben. Het stelt hen in staat om de gegevens ook voor andere doeleinden te gebruiken. Het gaat dan bijvoorbeeld om data die gebruikt wordt voor onderzoek.

Discussie

Naar aanleiding van de introductie van het conceptuele model en het doorlopen van de test hebben we aan de hand van een aantal vooraf bepaalde vragen gediscussieerd over het model en de huidige implementatie. Het uiteindelijke resultaat van

de testdoorloop en de concrete voorstellen ter verbetering zijn kort doorgenomen, maar niet uitgebreid besproken.

Gekeken naar *Data Driven Security Operations*, dan vindt *PC* dat het model een goede blik geeft op security operations. *PC* is bekend met andere invalshoeken op security monitoring en response, zoals de richtlijnen en quickscan van het Nationaal Cyber Security Centrum (NCSC), maar vindt dat het conceptuele model rondom data een nieuwe, verhelderende blik werpt op security operations. Ook noemt hij dat data en analyse daarvan steeds belangrijker wordt ten behoeve van informatiebeveiliging. Volgens hem zijn de zes verschillende onderdelen belangrijk. SOC staffing lijkt op het eerste gezicht een vreemde eend in de bijt, maar aan de hand van de vragen wordt wel duidelijk waarom dit wel een belangrijk facet is.

Op het eerste gezicht mist *PC* geen belangrijke aspecten in het model. Hij heeft wel het idee dat, ondanks dat informatie- en datadeling wel onderdeel is van het facet Threat Intelligence, dit onderbelicht blijft. Dit aspect vindt *PC* erg belangrijk, en daar had meer nadruk op gelegd kunnen worden. *PC* voegt daar aan toe dat er sprake is van *vetting* van personen: op basis van iemands naam en diens kunnen kan deze verwelkomd of juist uitgesloten worden van bepaalde gemeenschappen voor het delen van gegevens. Het is zelfs voorgekomen dat een bedrijf werd uitgesloten van deelname, totdat een vertrouwd persoon bij het bedrijf kwam werken.

De uitwerking van het model in de grafische weergave vindt *PC* een goede weergave. Wat hem wel opvalt, is dat er in de implementatie geen mogelijkheid is om een 0 te scoren. Dat is volgens hem wel het geval bij andere, soortgelijke testen. *PC* vindt ook dat deze methode, inclusief vragen, goed gebruikt kan worden om vergelijkingen tussen verschillende organisaties of momentopnames op verschillende momenten te maken. Dit komt mede door een aantal zeer specifieke vragen waarover je even moet nadenken. Juist door erover na te moeten denken, komt je soms lager uit dan je in eerste instantie zou zeggen.

De manier van vraagstelling is volgens *PC* wel geschikt, maar de test moet wel voorgelegd worden aan de juiste persoon. Die moet zelf ook actief bezig zijn binnen security monitoring, en

goed weten wat er speelt. Bij enkele vragen is er extra uitleg gewenst: zonder mondelinge toevoeging zouden er hele verschillende antwoorden gegeven worden op basis van interpretatie.

Op grond van de vragen kan een organisatie preciezer vaststellen waar er dan mogelijk tekort geschoten wordt. *PC* zou een overzicht van de vragen met de bijbehorende score zeker op prijs stellen, aangezien daarmee de rapportage direct inzichtelijk maakt, wat er verbeterd kan worden. Dit zou ook voor een rapportage aan hoger management goed van pas komen: het plaatje laat op een hoog niveau zien waar de organisatie nu staat, maar de specifieke vragen geven daarnaast concreet aan waar actie ondernomen moet worden. Het model geeft een gemiddelde, maar er zullen punten zijn waarop je gemiddeld slecht presteert. *PC* stelt een bijlage met de vragen en de score voor, ook in het geval van subsets van vragen, maar dan moet wel duidelijk zijn op basis van welke subsets een bepaald antwoord gegeven wordt.

MSSP2

On March 8th, 2016 a full evaluation was performed with MSSP2. We had a meeting with two employees: Participant D (*PD*) and Participant E (*PE*). *PD* is the manager of the Security Operations Center (SOC) since mid-2015. He has several years of experience in managing security teams, such as the people working in a SOCs At MSSP2 he will use this knowledge to effectively control operations of the SOC analysts and penetration testers. Before starting at MSSP2, *PD* worked at Fox-IT for 12 years. *PE* has recently become the product manager of the SOC proposition at MSSP2. In his role he positions the SOC as an important spearhead to guarantee the continuity of business processes. Implementation of a Security Information and Event Management (SIEM) can be part of the proposition, but continuous monitoring of the IT environment by the experts at MSSP2 may be even more important. The latter is what *PE* tries to sell to potential customers through fitting business cases.

Introduction

MSSP2 is een Managed Security Services Provider (MSSP) gevestigd in de Randstad. Het bedrijf is begin 2015 opgericht als antwoord op de vraag naar de groeiende vraag naar het uitbesteden van security services. MSSP2 onderscheidt zich van andere, vergelijkbare dienstverleners door het inzetten van beveiligingsexperts die constant de IT omgeving van klanten in de gaten houden. Daarnaast staan er altijd specialisten klaar om te reageren op beveiligingsincidenten.

Het moederbedrijf van MSSP2 wilde haar klanten graag een totaaloplossing kunnen leveren. Deze totaaloplossing bestaat uit de implementatie, het beheer en de monitoring van de beveiligingstechnologieën gebaseerd op de wensen van de klant. Daartoe behoren onder andere anti-virus, Intrusion Prevention System (IPS), firewalls, email filtering, en SIEM, vulnerability scans en *sandboxing* technologie. Daarnaast biedt het moederbedrijf van MSSP2 *security as a service* aan via MSSP2. Op deze manier neemt zij alle verantwoordelijkheid op zich om de veiligheid van haar klanten te waarborgen. Het gaat onder andere om email filtering, vulnerability scans en managed anti-virus, waarbij het grote verschil hem vooral zit in het feit dat de organisatie niet hoeft te investeren in soft- of hardware.

De security monitoring dienst die MSSP2 levert, houdt in dat de omgeving van klanten continu en proactief gemonitord wordt op beveiligingsproblemen en incidenten. Het SOC wordt daarbij continu bemand door ten minste twee werknemers en is er geen sprake van piket diensten. Zodra er zich een beveiligingsincident voordoet, alarmeert en adviseert MSSP2 haar klant om adequate maatregelen te kunnen treffen. MSSP2 maakt daarbij gebruik van technologie die het mogelijk maakt om incidenten te detecteren die door op zichzelf staande beveiligingsoplosingen gemist worden.

MSSP2 implementeert en maakt gebruik van verschillende producten om haar diensten te kunnen verlenen. In 60 tot 70% van de gevallen wordt er gebruik gemaakt van producten van Intel Security, omdat het geheel aan producten dat zij aanbiedt, een volledige dekking van de IT omgeving van klanten kan bieden. Daarnaast wordt er gebruik gemaakt van FireEye, RedSocks en Fortinet producten. Binnen het SOC wordt er gewerkt

met McAfee ESM, dat tegenwoordig onder Intel Security valt, als SIEM, welke aangevuld wordt met AlienVault. De SIEM systemen worden gebruikt voor het detecteren en analyseren van misbruik of anderzins afwijkend gedrag. Dit laatste wordt gerealiseerd door het implementeren van vooraf gedefinieerde use cases.

Een ontwikkeling die *PD* ziet is de verschuiving naar *active response*: niet enkel het blokkeren van bijvoorbeeld uitgaande verbindingen met behulp van de firewall, maar ook het doorzoeken van endpoints op tekenen van hetzelfde gedrag. *PE* geeft aan dat er momenteel nog geen duidelijke Key Performance Indicators (KPIs) zijn voor security monitoring. Dit komt voort uit het feit dat technologie geen garanties kan geven over wat en hoeveel er gedetecteerd wordt. Wel kan technologie garanderen dat er binnen een bepaalde tijd gereageerd kan worden. Dat is ook hoe de dienstverlening van het SOC is opgezet: er worden harde cijfers gegeven die betrekking hebben op de reactietijd. Volledige bescherming bestaat niet: de verwachtingen van de eindklant worden daarop bijgesteld.

Validation

The validation of the research was also part of the meeting on March 8th. First we did a quick walk-through of the conceptual model during which each of the six facets was discussed to quite some extent. After the introduction some pointers for performing the test were introduced, after which the test was performed by *PD*. All of the remarks mentioned during and after the test have been included in the following summary.

Het gesprek van 8 maart 2016 is samengevat in de volgende secties. Het gaat om alle inhoud besproken voor, tijdens en na het doorlopen van de test.

Presentatie conceptueel model

Het eerste onderdeel was de presentatie van het model voor *Data Driven Security Operations*. Aan de hand van een presentatie met de kernpunten van elk van de facetten, werd het model mondeling toegelicht, waarbij er door *PD* en *PE* uitgebreid ge-

bruik gemaakt werd om opmerkingen te plaatsen en vragen te stellen.

De eerste blik op het conceptuele model zorgt niet voor hele grote vragen of veel onduidelijkheid. Wel vraagt *PD* zich af of er overlap bestaat tussen de facetten *Detection Methods* en *SOC Infrastructure*. Ook levert de afbeelding geen gedeeld begrip van elk van de facetten zonder uitleg over wat elk van de facetten inhoudt. Bij het onderdeel visualisatie, noemt *PD* dat er binnen MSSP2 ook gewerkt wordt aan het inrichten van nieuwe dashboards.

De indeling van *Threat Intelligence* op strategisch, tactisch en operationeel spreekt *PD* aan. Hij noemt dat op strategisch niveau vooral gekeken moet worden naar trends die gaande zijn, zoals het beveiligen van Industrial Control Systems (ICSs) (red: Industriële Controlesystemen) en het Internet-of-Things (IoT), en dan met name op de langere termijn. *PE* noemt dat er nogal een verschil zit tussen (jaar)rapportages. Een voorbeeld dat hij geeft is het Data Breach Investigations Report (DBIR) van Verizon, dat meer gericht is op kwantitatief ingestelde mensen in vergelijking met het Security Report 2016 van Mnemonic[10], terwijl deze wel allebei aangemerkt kunnen worden als *Threat Intelligence* op strategisch (en tactisch) niveau.

Het aspect van informatie delen zou volgens *PD* niet voor elke MSSP vanzelfsprekend zijn. De kennis over bepaalde dreigingen is een onderdeel van het differentiërend vermogen van een MSSP, en zou dus gezien kunnen worden als intellectueel eigendom wat je niet zomaar met iedereen deelt. Wel ziet hij voordelen van het kunnen delen van informatie, maar dan wel in een vorm waarin de informatie niet direct vrijgegeven wordt, bijvoorbeeld door het versleuteld vergelijken van indicators, en pas waarna daar reden voor is, meer informatie prijs te geven.

PD en *PE* zien voordelen in het toevoegen van contextuele gegevens aan bijvoorbeeld log data om *false positives* te verminderen en de analist meer informatie te geven. Ze geven daar wel bij aan dat er aan de ene kant complexiteit wordt toegevoegd aan het realiseren van een datastroom met context waar dat bij de

10 https://www.mnemonic.no/globalassets/security-report-2016/mnemonic_security_report_2016.pdf

analist weggenomen wordt. Aan de engineering kant van het verhaal kost dat tijd en voornamelijk geld, wat naar alle waarschijnlijkheid beter te realiseren is door een intern SOC dan door een MSSP. *PD* noemt als concreet voorbeeld de oplossingen van RedSocks. Standaard wordt een groot deel van de malware correct gedetecteerd, maar dit gaat gepaard met een groot aantal *false positives*. Het toevoegen van context, zoals reputatie data voor domeinen, het al dan niet aanwezig zijn van data in *post requests* en het downloaden van data geassocieerd met uitvoerbare bestanden, zou de basis-functionaliteit van deze oplossing sterk kunnen verbeteren. *PD* noemt dat er in dit soort gevallen ook actie wordt ondernemen richting de productpartners om verbetering te realiseren.

Het openstellen van, of de mogelijkheid bieden om aanpassingen te maken aan de statistische en/of *machine learning* modellen zal volgens *PD* een lastig punt zijn. Hij noemt dat je als MSSP steeds minder controle hebt over hoe een product precies werkt, en dat sommige vendors de garantie en/of support staken zodra er wel wijzigingen worden aangebracht. *PE* vraagt zich af in hoeverre je iets kan doen met of veranderen aan de statistische modellen voor de IT omgeving. Ook over zulke statistische modellen heb je als SOC weinig controle, en ze zijn afhankelijk van de IT omgeving. Wel kunnen ze gebruikt worden voor het visualiseren van verschillen, bijvoorbeeld over bepaalde tijdsspannen. Het is dan nog steeds niet mogelijk om (meteen) precies uit te leggen waarom een alert precies getriggerd is, maar het is wel inzichtelijk te maken. Ten slotte is *PE* nog erg benieuwd wat het feit dat detectie vooral *output driven* moet zijn, precies inhoudt.

De focus van MSSP2 zelf ligt voornamelijk op het bieden van managed services, waaronder security monitoring. Vanuit het moederbedrijf van MSSP2 wordt er echter wel gereageerd op incidenten die binnen kunnen komen via MSSP2 of andere wegen. *PD* zegt dat het real-time toegang hebben tot forensische data, zoals log data, *packet capture*, geheugen- en schijfimages vaak niet van toepassing is. In zulke gevallen is bijvoorbeeld de data al verdwenen, als er al loggegevens (centraal) opgeslagen werden met behulp van log management of SIEM oplossingen. Hij noemt de kosten en performance impact als belangrijkste redenen hiervoor. De gewenste data is tegenwoordig ook door bijvoorbeeld uitbesteding van de IT of het werken in de cloud

niet beschikbaar. De afwezigheid van deze data zorgt ervoor dat je als puntje bij paaltje komt, je geen volledig uitsluitsel kunt geven over een incident.

Toegang tot bijvoorbeeld *full packet capture*, bijvoorbeeld na het triggeren van een detectieregel of door *rolling capture*, zouden voor een analist de meest rijke bron van informatie zijn. *PE* noemt dat een project bij een eerdere werkgever, waarbij er gebruik gemaakt zou worden van *full packet capture*, uiteindelijk afgeschreven is, omdat de opslag van alle data te veel tijd en geld zou kosten.

PD legt uit dat het Incident Response Team (IRT) niet heel vaak in actie hoeft te komen, omdat in veel gevallen een incident al tijdens de initiële triage kan worden afgehandeld. Als het IRT wel in actie moet komen, is de eerste stap het scheiden van management en operationeel is, om zo te voorkomen dat het een te bureaucratisch proces wordt. Het IRT kan daarna werken aan het vinden van de oorzaak, onderzoeken van de impact en het remediëren van het incident.

PD noemt dat de voorbereiding een belangrijk onderdeel van het Incident Response (IR) proces is. Er zijn verschillende bedrijven die zich met dat aspect van IR bezig houden, waarbij onder andere *forensic readiness* komt kijken. *Forensic readiness* draait om de paraatheid van een organisatie met betrekking tot het reageren op en afhandelen van een beveiligingsincident. Dit kan ook inhouden dat een externe partij deze taak voor hen uitvoert, maar dan moet daar wel een proces voor ingericht zijn. Concrete voorbeelden van zaken die voorbereid kunnen worden zijn het bepalen van de kritieke systemen en welke data benodigd is in het geval van een incident. Ook het ervoor zorgen dat men daadwerkelijk over de juiste data beschikt is iets dat voorbereid kan worden.

PD beaamt dat het beheersen van adequate kennis op het gebied van informatiebeveiliging een must is, maar dat daarnaast de mindset van werknemers zeker een toegevoegde waarde heeft. Hoe warm loopt men voor de ontwikkelingen op security gebied en houdt een werknemer zich bijvoorbeeld ook buiten zijn werk bezig met (eigen) projecten? Hij noemt dat hij daar tijdens sollicitatieprocedures ook al naar kijkt.

Testdoorloop

Voordat met de doorloop van de test gestart werd, is uitgelegd hoe de test precies in elkaar steekt. Aangezien MSSP2 een MSSP is, is het uitgangspunt voor het beantwoorden van de vragen om te kijken naar de top 3 van klanten.

Aandachtspunten tijdens testdoorloop

- Voor een MSSP is het lastig in te schatten wat de kritieke bedrijfsprocessen zijn. De bijbehorende vraag is daardoor niet heel concreet te beantwoorden.
- De vraag over de compleetheid van het *threat model* is lastig te beantwoorden. Bij klanten waar MSSP2 een SIEM levert vormt dat de basis. Buiten SIEM is er een holistische aanpak, met bijvoorbeeld Intrusion Detection System (IDS), email gateways voor de detectie en SIEM is gericht op de bedrijfsprocessen. *PD* wil meer technische use cases in de SIEM i.p.v. de traditionele aanpak die op compliance gericht was.
- *PD* noemt dat het mooi zou zijn om met de *attack graph* de *kill chain* heel duidelijk inzichtelijk te maken voor de datastromen. De *kill chain* houdt voor *PD* in dat je inzichtelijk welke stappen er genomen zijn, en gaat volgens hem niet zozeer om de standaard die Lockheed ontwikkeld heeft.
- De verschillende vragen over inventories (hardware, software, configuraties) lopen door elkaar: zijn ze allemaal nodig?
- *Identity management* valt momenteel heel erg tegen.
- Bij het punt over de compleetheid van de dataverzameling noemt *PD* dat organisaties in sommige gevallen dit aspect wel in orde kunnen hebben, maar dat er alsnog sprake kan zijn van een incompleet beeld. Een MSSP heeft niet altijd de volledige controle over wat er bij een klant staat. Als de data wel bij de MSSP binnenkomt, kan die bijgesteld worden indien nodig, maar als specifieke functionaliteit op apparatuur uitgeschakeld staat, dan krijgt een MSSP daar geen inzicht in. *"bullshit in is bullshit out"*.
- Af en toe wordt er Threat Intelligence gecreeërd en gedeeld met de industrie.

- Het is *PD* niet meteen duidelijk hoe Threat Intelligence gebruikt kan worden om het risicoprofiel van een organisatie beter te snappen. *PE* voegt daaraan toe dat Threat Intelligence gebruikt kan worden om een beter beeld te krijgen van wat er speelt om zo de risico's beter te begrijpen.

- Gebruik en begrip van Threat Intelligence gaan hand in hand: meer gebruik leidt tot meer begrip en andersom.

- *PD* vindt dat er op grote lijnen geen hele grote verschillen zijn tussen organisaties. Threat Intelligence gericht op specifieke *verticals* staat volgens hem nog in de kinderschoenen.

- Met reactief onderzoek heeft MSSP2 niet vaak te maken. 90% is gerelateerd aan endpoints of beleidsmatig. Er hoeft dan geen forensische analyse uitgevoerd te worden; de onderste steen hoeft niet boven. Dat is wel het geval bij een *breach*, maar die vinden niet vaak plaats. Het SOC houdt zich echt bezig met het reactief monitoren en voor het onderzoeken van incidenten is dan eigenlijk altijd een apart team beschikbaar. De focus van MSSP2 ligt dus op *incident handling* en *triage*.

- De vraag naar *formeel* ingerichte processen voor zelfstudie en -ontwikkeling vindt *PD* een op zijn zachtst gezegd irritante vraag.

- De combinatievraag over OSINT en *reverse engineering* is niet makkelijk om te beantwoorden, omdat het gaat om twee verschillende dingen.

- Het concept van een *linked data model* is *PD* niet meteen duidelijk. Na een korte uitleg en voorbeeld voegt hij toe dat het wel te doen is om zo'n model te construeren, maar dat het nog niet te visualiseren is.

- Het inschatten van de schaalbaarheid van de SOC infrastructuur is lastig omdat er met verschillende producten gewerkt wordt. Deze moeten met elkaar samenwerken en tegelijkertijd inzichtelijk zijn. Er is op dit moment nog geen master dashboard waarin verschillende dashboards voor verschillende klanten samengebracht worden.

- Een van de zaken waaraan *PD* en zijn collega's aan gaan werken is het maken van eigen dashboards. Daarbij gaan ze gebruikmaken van Logstash en Kibana om onder andere de data van de eigen honeypots te visualiseren. Het putten uit de database van McAfee ESM wordt momenteel nog wel achterwege gelaten omwille van zaken omtrent technische support op het product.

Discussie

Na het doorlopen van de test werd de eindrapportage gegenereerd met een korte uitleg over hoe de score tot stand is gekomen. Het eerste wat *PE* opmerkt bij het zien van het gescoorde model is dat de score voor *Response & Investigation* hem niet verbaast, aangezien dat geen kernonderdeel is van de diensten van MSSP2 Daarna volgde uitleg over de onderdelen van het rapport. De doelstelling bleek bij vijf van de zes facetten hoger te liggen dan het huidige niveau op basis van de antwoorden op de vragen. Op het facet *Response & Investigation* scoort MSSP2 momenteel hoger dan de doelstelling.

Op de vraag van *PD* naar wat er gebeurt met de ingeschatte waarden word de afbeelding met de desbetreffende waarden getoond. Daarbij wordt uitgelegd dat deze afbeelding geen onderdeel is van het rapport, maar wel dient als een deel van de evaluatie van het onderzoek om vast te stellen in welke mate de gestelde vragen overeen komen met de opgebouwde verwachting(en). Het valt direct op dat de verwachtingswaarde voor *Response & Investigation* slecht aansluit bij de berekende waarde. Een andere uitschieter is *Detection Methods*, waarbij *PD* noemt dat hij deze expres iets hoger heeft ingevuld.

PD vindt dat het model een basis is die inzichtelijk maakt waar je rekening mee moet houden. De implementatie draagt ook bij aan het krijgen van een eerste beeld. De manier van vraagstelling en de interpretatie van de vragen maken dit beeld minder concreet. Er zou een verdiepingsslag over de vragen moeten gehaald moeten worden om te bepalen wat precies van toepassing is op een SOC en wat niet.

Het benoemen van verschillende voorstellen om de doelstelling te behalen spreekt zowel *PE* als *PD* aan. Wel benoemt *PD* dat bedrijven juist ook kijken naar commerciële producten om zelf

minder tijd en geld kwijt te zijn aan de installatie, onderhoud en ontwikkeling, dus dat open source niet voor elk bedrijf een reële optie is. *PE* noemt dat de uiteindelijke afweging tussen commerciële producten en open source oplossingen uiteindelijk een afweging is tussen *Capital Expenditures (CAPEX)* en *Operational Expenditures (OPEX)*; uiteindelijk zou je evenveel kwijt kunnen zijn ongeacht de keuze. Waar volgens *PD* in de implementatie beter rekening mee gehouden kan worden is het geaccepteerde risico.

B FACET DESCRIPTIONS

This appendix contains the full descriptions of each of the facets of the conceptual model. These are extended versions of the content as presented in Section 6.1.1, but do not contain references to (scientific) literature. The contents are presented here is also a part of the final report that organizations received after the assessment as constructed in Section 6.2.

SITUATIONAL AWARENESS

The most important thing a Security Operations Center (SOC) has to accomplish is gaining full Situational Awareness. Situational Awareness is defined as the continuous extraction and integration of environmental information, in order to form a coherent mental picture of the current situation. Current approaches to acquiring Situational Awareness include vulnerability analysis through attack graphs, intrusion detection and correlation, attack trend analysis, causality analysis and forensics, taint and information flow analysis, damage assessment and intrusion response. The perception of the current situation, on the strategic, operational as well as tactical level, can then be used to anticipate future events more readily. Security Operations Centers (SOCs) will not be able to adequately perform their functions when they do not fully understand their constituency's IT environment.

Attaining strategic Situational Awareness is important for making the organization aware of their current security status and risk profile. Reporting on the current risk status of the organization is one of the most important activities here. Decision makers should be informed about the risks the organization is susceptible to at the current moment, but also on the long-term, to allow them to incorporate the information into their decisions. Prerequisites for making this work are a complete understanding of the organization's business processes, their criticality and the risks involved.

At the operational level, SOCs have to align the threats and risks the organization is currently facing with their mission to protect the constituency, appropriately prioritizing for the most critical ones. Constructing a complete threat model and vulnerability analysis of the constituency and its (critical) business processes are key elements to attain operational Situational Awareness. On a more practical level, this encompasses the creation of attack graphs that are mapped to the constituency's IT environment to analyze the consequences of both hypothetical and simulated attacks and intrusions. The attack graphs can also be used to simulate malicious events in order to assess the level of protection provided by preventive, detective and corrective controls.

In order to realize Situational Awareness at the tactical level, the SOC needs to have complete visibility into and understanding of the constituency's IT environment. One of the requirements is to have access to databases containing the configuration, patch status and vulnerabilities of all systems, applications, services and identities. In addition to having a single up-to-date view of the IT environment at a given time, the SOC also needs data about what is currently happening, which comes in many forms, such as network traffic (e.g. full packet capture or rolling captures) and log data from hosts, services, applications and networking gear. All of this data needs to be integrated and analyzed in real-time to comprehend the current situation. Having the ability to quickly zoom in on specific assets or events from a high macroscopic level view is an absolute must for the SOC, because this will ultimately result in Situational Awareness at microscopic level.

Information visualization and appropriate methods for humans to interact with the data and information are the most important aspects of Situational Awareness on strategical, operational as well as the tactical level. As long as machines are not intelligent enough to protect themselves it is up to a human analyst to comprehend and project the current situation. The ability to create and analyze visualizations of the entire IT environment helps human operators to more quickly and completely comprehend the current situation. Effective dashboards provide analysts with overviews of the current situation supplying them with information extracted and enriched with context, including network statistics and event summaries. On the strategical level, the ability to create reports containing the right information to describe the current status is a necessity for sup-

porting the best decisions to be made. Creation and analysis of attack graphs is important on the tactical level. A model of the IT infrastructure showing interdependencies between systems that are augmented with security-relevant details can help an analyst get intuitive insights about the IT infrastructure and allows him to spot threats more quickly. Having the ability and flexibility to create new and improve visualizations is important on all three levels.

THREAT INTELLIGENCE

The second facet for *Data Driven Security Operations* is Threat Intelligence. It can be described as containing knowledge based on evidence, including context, Indicators of Compromise (IOCs), techniques, mechanisms and implications about existing or emerging threats. Threat gce is diverse and can be leveraged on strategic, operational as well as on the tactical level.

Strategic Threat Intelligence usually consists of high level descriptions of current threats and is distributed via reports or discussed during meetings. It is typically consumed by members of the C-suite or people that report to the board members and is focused on business and strategic guidance. By analysis of the reports a better understanding of current and emerging threats and risks can be gained which in turn allows for better decisions to be made with regards to managing these. Strategic Threat Intelligence can also surface risks the organization was not yet aware of.

On the operational level Threat Intelligence gives insights into current and impending attacks against the organization. This type of intelligence can be hard to realize within a Security Operations Center (SOC), because of limited insight into what specific threat actors are going to do. This limited insight is caused by having no access to the communication channels and infrastructure used by adversaries. In some cases operational Threat Intelligence can be produced from publicly available sources by performing Open Source Intelligence (OSINT) or by linking real-world events to (upcoming) digital events, for example. Nation states and Internet Service Providers (ISPs) have more control over the communication channels and infrastructure used by adversaries, which may result in obtaining a higher level of operational Threat Intelligence. Operational

Threat Intelligence is typically consumed by the higher layers of security staff, such as security and SOC managers.

On the tactical level, Threat Intelligence can be described as any technical details about specific threats, during both short-term as well as long-term time frames. Within the short-term time frame this includes the usage of specific instances of malware (e.g. MD5 hashes), domain names and IP addresses by threat actors. Organizations can also focus on longer-term efforts by adversaries, such as what Tactics, Techniques and Procedures (TTP) threat actors or groups are currently using. This information can be used to learn about how threat actors are likely going to perform an attack against them, which can then be used to assess the currently active controls and surface potential shortcomings. Tactical Threat Intelligence is increasingly being made available and distributed through standardized formats, such as Structured Threat Information eXpression (STIX) and Trusted Automated eXchange of Indicator Information (TAXII). Using only a single source of Threat Intelligence is not likely going to get great results, because several distinct feeds show minimal overlap in indicators. Combining several sources and analyzing them together before consumption is commonly performed in so-called Threat Intelligence Platforms, which allow for the collection, analysis, consumption, production and sharing of Threat Intelligence. Organizations should, after some extent of (manual) testing and evaluation, automate the usage of short-term tactical Threat Intelligence to a full or large extent, because the volumes of data will become unmanageable by human operators. This way the currently effective controls, such as firewalls, intrusion detection systems and endpoint protection systems, can be updated as soon as possible. Especially on the tactical level it is important to constantly monitor the effectiveness of Threat Intelligence usage, because it's not hard for an attacker to create new malware which can't be identified by existing signatures, for example. The most important aspect of good tactical Threat Intelligence is that it should be timely: it should be based on recent and current attacking campaigns. Besides that, older Threat Intelligence can give insights into historical threats: organizations can search their logs for the IOCs to uncover (unknown) incidents that happened in the past. Tactical Threat Intelligence is usually consumed by the people responsible for the IT infrastructure, including system administrators and security staff.

Threat Intelligence can prove really valuable for an organization, be it on strategic, operational or tactical level. Creating an effective Threat Intelligence program requires well-defined processes for requirements elicitation, data collection and analysis, evaluation and possibly production and sharing. Asking the right questions to set the requirements for the Threat Intelligence is one of the most important steps, which is followed by evaluation to eventually get better at using Threat Intelligence.

DETECTION METHODS

Detection Methods have been an important part of the day to day operations in the Security Operations Center (SOC). The SOC has always relied on data to function, such as extracting the strings from a file to determine whether it's malware or not and inspecting the traffic flows between endpoints. Additionally, certain indicators that can trigger an actual detection have to be present, which include signatures and normal behavioral models. Configuring intrusion detection systems is not particularly hard, but maintaining it and not over-configuring is are in fact hard to do, and may result in far too many false positives or not detecting anything at all.

Modern SOCs continue to deploy so-called Security Information and Event Management (SIEM) systems at the core of their infrastructure. The SIEM proved a valuable asset for SOCs, because it allowed the collection and integration of the data generated by many types of security systems, including firewalls and intrusion detection systems. Precisely crafted rule sets would then comb through the data to uncover threats hiding in the IT environment. This could in principle reduce the amount of alerts, but in practice this is not always the case, and SOCs still face an overload of alerts to analyze.

The central security data storage capability is certainly something that SOCs could continue to enforce, but detection methods currently employed are not likely to continue providing value for ever. Administration of the rule sets in place is a tedious task which does not offer a guarantee for catching malicious activity, which is why SOCs will have to improve their current practices. In a data driven approach to security, the aggregated data is supposed to be richer than in the usual SOC. Ideally, the data is contextualized, which means it has been enhanced for the security operator to analyze. The data should

provide a more complete view of the organization by having information about the organization embedded in it.

Making the current detection methods use the additional contextual information can result in less false positives, because more information is available. The creation of behavioral baselines for various entries, such as hosts, applications and users, becomes a possibility when data is collected and stored during extended periods. These profiles can then be used by the SOC to operationalize anomaly detection in a specific business context, but this demands specific knowledge in the fields of statistics and machine learning in addition to security monitoring.

Proper management of detection methods has always been an important aspect of successfully using them. This was (and still is) the case for intrusion detection systems, but becomes even more important in a *data driven security* setting. Processes should be in place to monitor currently active detection methods for performance, e.g. efficiency and efficacy and track those over time. Another important aspect of the Detection Methods facet is that methods should be output driven and be described clearly upfront. This is a process that starts with the question what to detect, how to detect it and determining what data is required to do so. These use cases can then be implemented in a system of choice, such as a SIEM or a different platform.

RESPONSE & INVESTIGATION

When a possible intrusion has been detected, the Security Operations Center (SOC) can escalate the event. The security event then becomes an incident: a violation or imminent threat of violation of computer security policies, acceptable use policies or standard security practices, which has to be handled. What is considered a security incident within a specific organization depends on its business processes, but common events that may trigger an incident include unauthorized system access, malware infections and data loss. The goal of incident response is to handle the security breach in such a way that limits the damage and reduces the time to recovery and the costs involved. As soon as a security incident is detected, it has to be analyzed to identify its root cause in order to resolve the incident and remediate it. Not every SOC performs incident response itself, but it remains a core part of security operations within an organization. In cases where the SOC is not responsible itself for

incident handling, external parties can be contacted to handle the incident.

Irrespective of which party is responsible for the incident handling process, having access to the right data is an absolute must within incident response. The data is necessary in order to investigate the incident, find out what happened, who or what entities were involved, assess the impact of the incident, how to recover from the incident and to actually recover from the incident. During the incident response process every second counts, so having access to the data and the right fidelity of data is crucial.

When the SOC has access to a centralized data storage infrastructure this greatly improves the incident response and investigation process. In the case that the SOC does not have access to the data directly it should be able to obtain the data on short notice. Having remote access to and control over all endpoints can prove invaluable in such a case, because analysts can start their investigation from that point. One example is that the SOC (or an automated system) can start a live packet capture or memory of disk images can be acquired instantly. Methods to reconstruct objects and sessions can provide analysts with an even more complete view of what happened at the time of the incident and can help identifying the root cause.

Investigation into incidents is mostly initiated only after an incident response process has been started and when deemed necessary: reactive investigation. We propose its logical counterpart, proactive security investigation (also popularly called hunting), to become a core part of the SOCs functions. Proactive investigation should be used to uncover threats in the security data that is already available. This method can be used to extract intelligence and create new detection methods, so that the organization does not fall prey to the same kind of threat in the future. Analysts need an interface to all of the available data resulting in unlimited flexibility during their analysis which also performs fast. This way they can take advantage of tools for visualization they already know about and can concentrate on the hunt.

In a mature *data driven security* strategy, we foresee the emergence of the continuous security response process. The continuous security response process consolidates the prevention of, detection of and response to threats and security events. It is the culmination of the *data driven security* strategy, in which data is transformed into information, then intelligence and even-

tually applied in practice to improve all of the SOC functions. Some examples of the process include the investigation of incidents resulting in new detection methods and increasing the visibility in areas where it fell short before the incident.

The main tenet of the incident response and investigation facet is that the processes for data collection, analysis and incident remediation are well-defined and repeatable. This allows the SOC to become nimble in an ever-evolving threat landscape and complex IT environment.

SOC STAFF

Hiring the right people for security operations has always been a hard task. A diverse skill set is necessary, including both theoretical and practical knowledge about network devices, network security analysis, network protocols, application security, engineering and software and tools in all of these fields. Depending on which functions the Security Operations Center (SOC) delivers, more specialist knowledge in reverse engineering, malware analysis and digital forensics may also be necessary to have available. Another human aspect important for success is the mindset of people: security monitoring and incident response are fields that will likely never be developed to their full extent. Security operators will have to be passionate and curious about their job and should continuously develop their skills and knowledge to stay relevant.

In addition to traditional functions within the SOC, several new ones can be introduced in a *data driven security* strategy. When the SOC wants to improve its activities around Threat Intelligence, security operators may need a deep understanding of Open Source Intelligence (OSINT) or structured analytic processes for creating real intelligence, depending on what level (strategic, tactical, operational) the SOC wants to improve. If a SOC wants to deploy new types of detection methods, such as statistical models or machine learning based approaches, security operators having skills related to these are an absolute necessity for a successful implementation. Setting up a specialist hunting team or individual requires them having practical knowledge in the field of programming, visualization and data science.

Besides SOC staff having certain skills, the SOC should also invest in the development of its members. Continuously hav-

ing training and education opportunities available will keep the operators up-to-date and not doze off. This is also supported by having clearly defined career paths available. Another aspect for success is creating a well-oiled team of operators. Having means available for knowledge sharing and collaboration, which can include predefined work flows and supporting technologies, will greatly improve the capabilities of the SOC.

SOC INFRASTRUCTURE

All of the facets of *Data Driven Security Operations* described before increasingly depend on the right types of technology and infrastructure to accomplish. Another aspect that all of the facets have in common is that they all involve data. To attain Situational Awareness, data has to be integrated and visualized in meaningful ways and at different levels of detail and abstraction. Threat Intelligence requires data analysis skills and ways to integrate it with the current Security Operations Center (SOC) infrastructure in order to make use of it. Current Detection Methods can be improved or new ones can be developed based on certain data enhancements and contextualization of the data. In order to respond to an incident in a timely manner, the right data and level of detail has to be available to start the investigation. When looking at continuous response and hunting, it is even more important to provide interfaces to access all of the data that is available. All of these also increase the need for the right people to be present in the SOC team, because adopting a data driven approach to security introduces new and previously less relevant requirements.

Besides supporting all of the functional requirements set by the SOC, the SOC infrastructure also needs to fulfill non-functional requirements. The first ones include requirements with respect to performance, such as the storage and processing capacity and how fast the data can be processed. It is also evident that no single technology can address all of the threats organizations face today, so the infrastructure should be extensible and foster compatibility. Another aspect that certainly shouldn't be excluded is the security of the SOC infrastructure itself. Centralized security data collection makes the storage systems attractive to adversaries which demands controls to be in place to maintain the confidentiality, integrity and availability of the data and the systems relying making use of the data.

C | MODEL SYNOPSIS

This appendix summarizes the essential parts of our model for *Data Driven Security Operations*. It's essentially a gist of the model's facets.

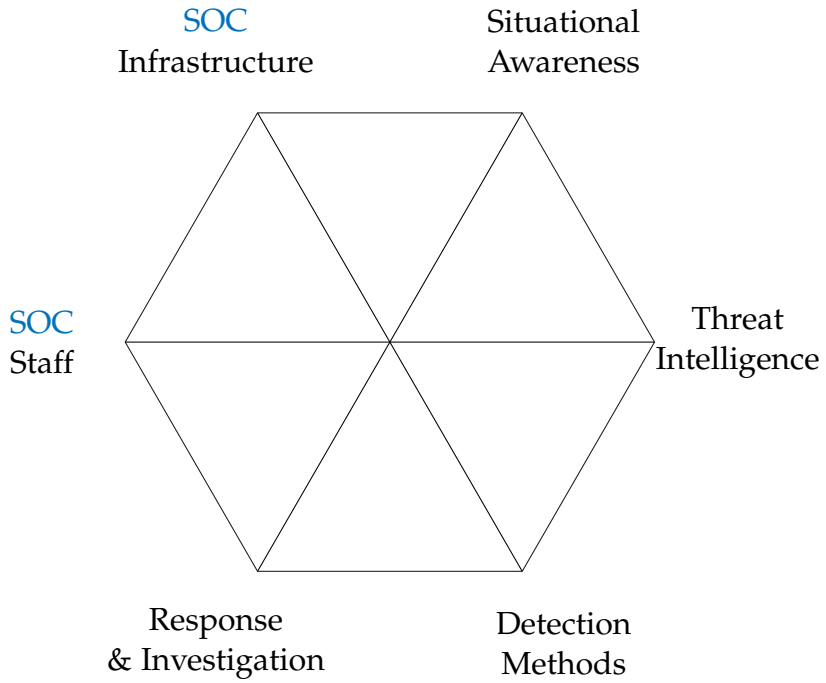


Figure C.1: Conceptual model for *Data Driven Security Operations*.

SITUATIONAL AWARENESS

- *Perception* and *comprehension* of the current IT environment and *projection* into the future.
- Continuous extraction and integration of environmental data.
- Has to be attained at *strategic*, *operational* and *tactical* levels.
- Abstraction on macroscopic as well as microscopic level.
- Effective visualization of the IT environment in both scope and time are important.

- Think graphs, not lists.

THREAT INTELLIGENCE

- Can be applied on *strategic*, *operational* and *tactical* level, depending on requirements and maturity of the Security Operations Center (SOC).
- Should be used to improve decision making, learn about Tactics, Techniques and Procedures (TTP) and increase context.
- Employing tactical Threat Intelligence should be highly automated.
- Well-defined processes for requirements elicitation, TI collection and analysis, evaluation, production and sharing should be in place.

DETECTION METHODS

- Improve current detection methods by contextualizing data.
- Continuously monitor and evaluate effectiveness and efficacy of detection methods.
- Aim for *output driven* detection: define a problem, construct (data) requirements, develop or improve, test and implement methods.
- Make use of *open* statistical and behavioral anomaly detection: aim to know what's going on under the hood.

RESPONSE & INVESTIGATION

- Aim for increased control over network devices and endpoints.
- Attain real-time (ad-hoc or continuous) access to forensic data.
- Move towards continuous response and proactive investigation into the IT environment, to potentially uncover undetected threats, shed light on the IT environment and improve detection.

- Data is only the start of the story. It has to processed to create information, analyzed to instill knowledge and applied to improve Security Operations Center (SOC) functions and gain wisdom.

SOC STAFF

- Appropriate skill set is important, but the right mindset might be of even bigger importance.
- Instill a culture of continuous development and learning.
- In addition to security and networking knowledge and skills, the Security Operations Center (SOC) needs intelligence, engineering, machine learning, statistics and data visualization knowledge and skills.
- Take advantage of analytical models to structure tasks and thinking.
- Improve team effectiveness through collaboration, predefined workflows and increased automation.
- Operationalize adequate security metrics to measure the effectiveness of the SOC.

SOC INFRASTRUCTURE

- All of the Security Operations Center (SOC) capabilities offered should be supported adequately by technology.
- Large scale data storage, processing, analysis and visualization forms the corner stone of *Data Driven Security Operations*.
- Due to changing demands and threat landscape, the scalability, performance, compatibility and extensibility of the SOC infrastructure are the main tenets to take into account.
- The SOC enclave should be operated securely itself.

D ASSESSMENT

This appendix contains all of the substantive questions used to calculate the mean score. Five different Likert ratings were used in total, which are shown in Table D.1.

Table D.1: Likert ratings used in the assessment

	Rating				
ID	1	2	3	4	5
A	never	rarely	sometimes	regularly	constantly
B	very poor	poor	acceptable	good	very good
C	non-existent	poor	acceptable	good	completely accurate
D	very low	low	moderate	high	very high
E	none	a bit	some	quite a bit	completely

Table D.2: Questions for Situational Awareness facet

ID	Question	Subset	Scale
SA1	To what extent has the criticality of business processes been defined?	I	E
SA2	To what extent is reported on the current risk profile of the organization?	I	E
SA3	Rate the quality and usefulness of reporting.	III	B
SA4	To what extent has the criticality of business processes been mapped to the (supporting) IT environment?	I	E
SA5	Rate the understanding of most vulnerable elements within the organization.	I	E
SA6	Rate the completeness and understanding of the threat model.	I	B
SA7	Are attack graphs constructed and evaluated?	III	A
SA8	Are 'what-if' scenarios being executed and evaluated?	IV	A
SA9	Rate the completeness of inventory with regards to hardware.	II	E
SA10	Rate the completeness of inventory with regards to software.	II	E
SA11	Rate the completeness of inventory with regards to identities.	II	E
SA12	How complete and accurate is the network topology map at any given time?	II	C
SA13	How complete and accurate is the configuration management databaes at any given time?	II	C
SA14	How complete and accurate is the vulnerability management databaes at any given time?	II	C
SA15	Rate the completeness of data collection from the IT environment.	III	B
SA16	Rate the ability to create a single, real-time view of the IT environment.	III	B
SA17	Rate the level of contextualization of the data being collected.	III	B
SA18	Rate the ability to quickly zoom in on specific entities from a high level view.	IV	B
SA19	Rate the ability to focus on a specific historical timeframe.	IV	D

Table D.3: Questions for Threat Intelligence facet

ID	Question	Subset	Scale
TI1	Rate the understanding and usage of Threat Intelligence (TI) on strategic level.	I, II	D
TI2	Rate the understanding and usage of TI on tactical level.	I, IV	D
TI3	Rate the understanding and usage of TI on operational level.	I, III	D
TI4	Rate the quality of the process for eliciting requirements, collecting, analyzing, producing and evaluating TI.	II, III, IV	B
TI5	Rate the extent to which TI is being used to better understand the current risk profile of the organization?	II	B
TI6	Rate the level of analysis performed on TI on strategic level.	II	B
TI7	To what extent is TI being used to gain an understanding of attacker modus operandi?	III	D
TI8	Rate how well the knowledge about attacker modus operandi is being related to the organization?	III	D
TI9	Rate the completeness of usage of TI across security operations.	IV	D
TI10	Rate the level of automation for the processes for collection and operationalization of TI.	IV	D
TI11	To what extent is TI being created and shared with the industry?	V	D

Table D.4: Questions for Detection Methods facet

ID	Question	Subset	Scale
DM1	Rate the level and quality of detection methods currently in use.	I	D
DM2	Rate the level of usage of anomaly detection.	I	D
DM3	To what extent are existing detection methods being improved?	II	A
DM4	To what extent are new detection methods developed, tested and operationalized effectively?	II	A
DM5	To what extent are detection methods being reviewed for effectiveness?	II	A
DM6	Rate the usage of deception techniques with respect to detection.	III	D

Table D.5: Questions for Response & Investigation facet

ID	Question	Subset	Scale
RI1	Rate the level of continuity in Incident Response (IR).	I	D
RI2	Rate the level of automation and repeatability in IR.	I	D
RI3	How regular is reactive investigation performed?	I	A
RI4	How regular is proactive investigation performed?	II	A
RI5	To what extent are processes well-defined and repeatable in reactive investigation?	II	D
RI6	To what extent are processes well-defined and repeatable in proactive investigation?	II	D
RI7	To what extent are new methods and procedures developed and evaluated for reactive and proactive investigation?	I, II	A

Table D.6: Questions for SOC Staff facet

ID	Question	Subset	Scale
SS1	To what extent have career paths of SOC staff been defined formally?	I	D
SS2	To what extent is (self)education supported and defined formally?	I	D
SS3	To what extent are training and certification supported and defined formally?	I	D
SS4	To what extent is knowledge sharing supported and defined formally?	II	D
SS5	To what extent is collaboration supported and defined formally?	II	D
SS6	Rate the currently present combined knowledge in and or experience with Security Monitoring.	III	D
SS7	Rate the currently present combined knowledge in and or experience with Digital Forensics.	III	D
SS8	Rate the currently present combined knowledge in and or experience with Programming.	III	D
SS9	Rate the currently present combined knowledge in and or experience with Engineering.	III	D
SS10	Rate the currently present combined knowledge in and or experience with Statistics and/or Machine Learning.	III	D
SS11	Rate the currently present combined knowledge in and or experience with Open Source Intelligence (OSINT) and/or Reverse Engineering.	III	D
SS12	To what extent have workflows been standardized and defined formally?	II	D
SS13	Rate the quality of operational metrics defined for SOC success.	IV	D
SS14	Rate the ability to adequately address all low and medium critical rated alerts.	V	D
SS15	Rate the ability to adequately address all highly critical rated alerts.	V	D

Table D.7: Questions for SOC Infrastructure facet

ID	Question	Subset	Scale
SI1	Rate how well the SOC infrastructure currently supports security operations.	I	D
SI2	Rate the completeness of data collection.	II	D
SI3	Rate the level of data fusion that can be accomplished.	II	D
SI4	To what extent can a linked data model be developed and visualized?	III	D
SI5	Rate the compatibility and extensibility of the SOC infrastructure.	I	D
SI6	Rate for the performance of the SOC infrastructure.	IV	D
SI7	Rate the scalability of the SOC infrastructure.	IV	D
SI8	Rate the security of the SOC infrastructure.	V	D

E REPORT RECOMMENDATIONS

This appendix contains all of possible recommendations that the implementation can generate and which become part of a personalized report. In the report the order of facets can be different based on the absolute difference between the set target level and the scored level for the facet. Within a facet the order of the recommendations can also be different based on the described absolute difference in descending order and on their relation to a subset of questions in the questionnaire. Hyperlinks and references have been included for completeness. A complete example report can be found on https://hermanslatman.nl/ddsoqs/.

SITUATIONAL AWARENESS

There seems to be a mismatch between the business and security departments. In order to not only get a view of the IT environment, but also understand it deeply, one has to understand the business and what risks the business faces. These risks and threats should be modelled. After complete understanding, the current situation can be projected in the future: what will be the biggest risks and should we thus prioritize?

No up-to-date view of the current IT environment can be established. Consider performing internal network scans that identify all active hosts, services and applications. These scans can be extended by performing vulnerability and configuration assessments. Technologies like osquery[11] may prove usable not only in a security setting, but also during business operations, because of its live query capabilities on connected hosts.

Not enough visibility into the IT environment is realized by the current infrastructure. This can be due to the fact that the log sources are not configured properly or not enough of them have been configured. Telemetry from network devices, hosts, services and applications can all be important to get full visibility into the IT environment. All of the data has to be col-

11 https://osquery.io/

lected and integrated centrally in a highly performant manner allowing high volume data storage also. Some highly performant technologies for data processing and collection have been developed in the recent past, including Kafka[12], Logstash[13], Storm[14] and Spark[15]. Efficient data storage and processing can be accomplished with Accumulo[16], Cassandra[17], Gaffer[18], HBase[19] and GraphX[20]. Another category of new technologies are the ones for providing access to the aforementioned storage technologies, such as Hive[21] and Elasticsearch[22]. Many of these technologies have been combined into ROCK NSM[23] and Metron[24], an open source platform for security analytics and successor to OpenSOC[25].

Seeing what is going on within an IT environment is only the beginning. The next steps are to increase the understanding of what is going on and to project the current situation into the future. These activities are largely based on mental processes by individuals which can be improved by applying analytical models. Two examples include the Kill Chain[26] and the Diamond Model for Intrusion Analysis[27]. These models can be used both for after the fact analysis, such as finding out what steps were taken during an intrusion, as well as for reasoning about future intrusion and ways to mitigate those.

Adequate analysis of the IT environment should be performed on both macroscopic as well as microscopic levels. These two levels should be realized in both the space as well as time dimensions. Analysts should have the ability to see the entire environment as well as be able to look at a specific entity, such as a host or user. Furthermore they should be able to look at

12 https://kafka.apache.org/
13 https://www.elastic.co/products/logstash
14 https://storm.apache.org/
15 http://spark.apache.org/
16 http://accumulo.apache.org/
17 http://cassandra.apache.org/
18 https://github.com/GovernmentCommunicationsHeadquarters/Gaffer
19 http://hbase.apache.org/
20 https://amplab.cs.berkeley.edu/publication/graphx-grades/
21 http://hive.apache.org/
22 https://www.elastic.co/
23 http://rocknsm.io/
24 https://metron.incubator.apache.org/
25 https://opensoc.github.io/
26 http://www.lockheedmartin.com/content/dam/lockheed/data/corporate/documents/LM-White-Paper-Intel-Driven-Defense.pdf
27 https://www.threatconnect.com/wp-content/uploads/ThreatConnect-The-Diamond-Model-of-Intrusion-Analysis.pdf

what happened in the environment or what a specific entity did at a specific time or during specific (aggregated) timeframes. Several academic examples include Cauldron[28], CyGraph (built on the neo4j[29] graph database) and AMICA, which started as projects at MITRE[30].

THREAT INTELLIGENCE

There does not seem to be enough understanding or usage of Threat Intelligence amongst all of the ranks within the organization. These include the strategical, tactical as well as the operational ranks. A comprehensive whitepaper[31] published by MWR InfoSecurity might prove a viable start for improving the understanding of Threat Intelligence. Besides that, starting small is a good thing to do.

On strategic level, Threat Intelligence should give guidance to what (kind of) threats are currently worth watching more closely. The knowledge gained by analyzing reports should be applied in risk-based optimization of information security. The existence of predefined processes for requirements illicitation and intelligence analysis can help an organization with regards to Threat Intelligence. Usually the Threat Intelligence that is consumed at a strategic level is more high level, and describes several kinds of inter-related trends, such as specific threats and changes to the IT environment, such as the IOT and industrial systems getting connected to the Internet. Some examples of sources for strategic Threat Intelligence include annual (trend) reports coming from Mnemonic, Verizon and ENISA.

Translating the intelligence to tactical decisions is falling short. This can be improved by getting better at understanding Threat Intelligence and how it should be applied. Threat Intelligence can be used at the tactical level to acquire branch-specific knowledge about attackers tools, techniques and processes (TTPs) which can be used to implement preventive, detective and corrective controls that counter these. Sitting between the strategic and operational layers, tactical Threat Intelligence consists of both trends as well as more contextual in-

28 http://cyvisiontechnologies.com/section/Cauldron/9/
29 http://neo4j.com/
30 https://www.mitre.org/
31 https://www.mwrinfosecurity.com/assets/Whitepapers/Threat-Intelligence-Whitepaper.pdf

formation. Examples of sources for tactical Threat Intelligence include reports describing APTs and reports like the ones created by ENISA, DELL, IBM, Verizon and Symantec (Internet Security Report).

Using Threat Intelligence at the operational level currently falls short. Effective usage of Threat Intelligence demands careful requirements elicitation and evaluation. There are many free feeds available, such as the Critical Stack Intel Feed[32], which is optimized for consumption by Bro[33]. More specialized feeds are usually available via commercial packages. Additionally, there exist search engines specifically aimed at providing Threat Intelligence, such as Deepviz Threat Intel[34]. Many more tools for using Threat Intelligence within your IT environment can be found in the following GitHub[35] repository.

On the collection and aggregation side of things so-called Threat Intelligence Platforms (TIPs) are emerging. These allow for collection, aggregation and analysis of numerous different types of sources of intelligence. Some open source examples include CIF[36], its successor Bearded Avenger[37] and threat_note[38]. The solutions offered by Anomali[39] (formerly ThreatStream) are an example of commercial solutions.

Sharing of Threat Intelligence is also a big thing in the security industry now. There are numerous efforts for sharing, including X-Force Exhange (XFE)[40] and AlienVault's Open Threat Exchange[41]. Organizations can also deploy their own Threat Intelligence Platforms (TIPs). Examples include MISP[42] and IntelMQ[43].

32 https://intel.criticalstack.com/
33 https://www.bro.org/
34 https://intel.deepviz.com/recap_network.php
35 https://github.com/hslatman/awesome-threat-intelligence
36 http://csirtgadgets.org/collective-intelligence-framework
37 https://github.com/csirtgadgets/bearded-avenger
38 https://github.com/defpoint/threat_note
39 https://www.anomali.com/
40 https://exchange.xforce.ibmcloud.com/
41 https://www.alienvault.com/open-threat-exchange
42 http://circl.lu/services/misp-malware-information-sharing-platform/
43 https://www.enisa.europa.eu/activities/cert/support/incident-handling-automation

DETECTION METHODS

One of the problems that SOCs face today is an overload of events and alerts. Security operators have to quickly determine which events or alerts are important to either send alerts through for deeper analysis, or discard them as false positives. One of the ways to counter the overload is by improving current detection methods. The creation of more specific is an examples of this. Despite clear downsides, such as the fact that it's tedious and error-prone work to perform, this is still being done today.

Existing detection methods can also be improved by contextualizing events. This means that events get enriched with additional information that was not available before. Adding the hostname to an event is a form of basic enrichment, but adding more advanced data, such as geolocation, reputation and HR data is also possible. Adding the identity of a user to an event can, for example, be used to trigger specific alerts for that person, uncovering an insider threat. Adding reputation data to DNS requests can quickly surface lookups and connections to domains with a bad or no reputation. Even when these types of data are not used by the detection methods themselves, the manual triage of an alert can be sped up by having this information available to the analyst.

One of the things that the current security industry bets on heavily is machine learning, including a related concept called (User and) Entity Behavioral Analytics ([U]EBA) and some of its offshoots. This type of products promises a lot of innovative detection methods based on machine learning and artificial intelligence, but in most cases they do not, for obvious reasons, provide exact specific specifications of what goes on under the hood. The most important drivers of behavioral analytics solutions are the concept of establishing entities and the creation of baselines for these entities and relationships between them. The latter depends on the availability of formal graph models. Some examples of products include Niara[44], Exabeam[45] and various offerings by Securonix[46]. These solutions often also interface with existing security technology, such as SIEM systems. Besides that, [U]EBA is also increasingly being used to improve

44 http://www.niara.com/products/advanced-analytics/
45 http://www.exabeam.com/
46 http://www.niara.com/products/advanced-analytics/

existing solutions by extending those, like Splunk[47] (powered by Caspida[48]).

Detection Methods The quality of detection methods is an aspect of the model that may be hard to grasp at first. In short, better detection methods are easy to maintain, have low false positive and negative rates and can be executed at high speeds. Processes should be in place to monitor and evaluate the currently deployed detection methods to assess which are performing well and which can probably be deactivated. Which is true for whatever type of detection method, is that they should be output driven, which can be described as a three step process. First you have to determine the goal you want to realize and how to do so, then get access to the necessary data and then implement and evaluate the method. Sometimes the data is already available, but in other cases you will need to enrich existing data or get access to and integrate completely new data.

In the current security market there's a lot of buzz going on about machine learning and anomaly detection. Research into applying anomaly detection to information security cases, especially in the academic world, has been ample and a lot of successes have been accomplished in this area. These successes are not reflected in practice however, which has a number of reasons, including the absence of relevant and up-to-date data sets and a lack of adequate ground truth. The fact that errors in anomaly detection and machine learning models are a lot more expensive than in other applications, such as recommendation systems, also plays a large role. Organizations should strive for a no-nonsense approach to applying machine learning to detection of threats, which demands specific knowledge and skills in the field of computer science. Models should be based on ground truth data and should work with the right features. Constructing and evaluating these models takes a lot of time, skills and continuously updating with the environment. Some resources and examples of more rigourous approaches to applying machine learning to information security include MLSec[49], the MLSec Project[50] and the Stratosphere IPS project[51].

47 http://www.splunk.com/en_us/products/premium-solutions/user-behavior-analytics.html
48 http://www.splunk.com/en_us/investor-relations/acquisitions/caspida.html
49 http://www.mlsec.org/
50 http://www.mlsecproject.org/
51 https://stratosphereips.org/

Deception is a (re)emerging topic in the security industry. It basically involves organizations deploying deceptive techniques, such as honeypots and honeytokens to trap attackers. The idea is to catch malicious actors in their tracks before they get to the real assets the SOC protects. This information can also be used to extract Threat Intelligence.

The Honeynet Project[52] is a scientific effort to increase the effectiveness of data analysis approaches, development of unique security tools gathering data about attackers and malicious software they using honeypots. Another example is the Modern Honey Network[53] which consist of multiple sensors running Snort[54], Conpot[55] (an Industrial Control Systems honeypot) and Dionaea[56], which is a honeypot for malware. and centralized management.

RESPONSE & INVESTIGATION

Many SOCs experience a breach at some time: it can happen to any organization. A quick incident response process can be the difference between catching the attacker early in his tracks or a major data breach or other crisis. The quicker an analyst can start and perform his reactive investigation, the quicker the conclusions can be drawn with regards to possible damages done.

Starting a reactive forensic investigation demands relevant data to be available. This can be collected ad-hoc with tools like Crowd Response[57], FastIR Collector[58] or FTK Imager[59], but tools like these might require access to the hosts affected. Alternatives include osquery[60], GRR Rapid Response[61], MIG[62] and FIDO[63], which all provide remote and ad-hoc access to hosts and can also be automated to some extent. Most of these projects are backed by large organizations, like Facebook,

52 http://honeynet.org/
53 https://github.com/threatstream/mhn
54 https://www.snort.org/
55 http://conpot.org/
56 http://www.edgis-security.org/honeypot/dionaea/
57 http://www.crowdstrike.com/community-tools/
58 https://github.com/SekoiaLab/Fastir_Collector
59 http://accessdata.com/product-download/?/support/adownloads
60 https://osquery.io/
61 https://github.com/google/grr
62 http://mig.mozilla.org/
63 https://github.com/Netflix/Fido

Mozilla and Google. Many more incident response related tools can be found in the following GitHub[64] repository.

Proactive investigation, or hunting, can be defined as the collective name for any manual or machine-assisted techniques used to detect security incidents that evaded automated solutions, like alerting rules or behavioral baselines. The idea is to analyze already available data using for example statistical methods or machine learning to find traces of malicious activity and to create new methods for detecting these in the future. Effective hunting starts with creating hypotheses, investigation and discovery using adequate techniques and technology and feeding potentials findings back into the hunting loop and existing and new detection methods. New hypotheses can be formed as a result of previous hunts, existing threat and vulnerabilitty intelligence capabilities and based on the current situation.

One of the techniques that one can use to find malicious tracks is by performing linked data analysis. This basically revolves around creating a data model that links different entities, possibly originating from different sources, together into a single directed and annotated graph. Visualizations of such a graph can be used to quickly spot strange behavior. One example of a product that provides this functionality is Sqrrl[65], which is based on Apache Accumulo[66]. Other examples of graph analytics products aimed at security include specific offerings by Palantir[67] and MapR[68]. The latter is one of several providers of enterprise-grade Hadoop[69] distributions and offers GraphX[70] running on top of Spark[71] out of the box. Another example is Forcepoint's (Raytheon) Sureview Analytics[72], which uses a federated approach to linking data from various sources and can be used to perform, a.o. link and temporal analysis.

64 https://github.com/meirwah/awesome-incident-response
65 http://sqrrl.com/
66 https://accumulo.apache.org/
67 https://www.palantir.com/
68 https://www.mapr.com/
69 https://hadoop.apache.org/
70 https://spark.apache.org/graphx/
71 https://spark.apache.org/
72 https://www.forcepoint.com/product/advanced-analytics/sureview-analytics

SOC STAFF

So-called alert fatigue is a common problem in security monitoring. SOCs are facing an increased amount of events to analyze and work through, of which many are false alarms (or: false positives). Investing in more security analysts is a solution, but may not be sufficient in the future, when interconnectivity of business processes increases even more. Performing risk-based prioritization might proof to be of more use. This can be accomplished by contextualizing events with business-specific information and fusing data for automatic analysis. Also critically reviewing what events and alerting rules you really need and which not, in order to reduce the number of sources and thus noise, might proof a valuable approach.

Continually sharpening the skills of analysts is a key element in a well-running SOC. The benefits are ample, including your analysts not getting bored and getting smarter at what they do. Making them learn new skills or focusing on specific specializations will unlock opportunities for the SOC as a whole. Having a clear career path shaped up front is a powerful incentive to work towards success for an individual analyst, but can also have benefits for the team as a whole, such as better staff retention in the competitive field of information security.

Having an overview of how well the SOC operates is important to measure its success. Many SOCs still operate using quantitative measures like the number of alerts or incidents handled and the number of (still) vulnerable machines. These kinds of security metrics do have in a current SOC, but should be extended with more qualitative approaches for measuring success. These may include the means of time measurements, such as the Mean Time To Detect (MTTD), Mean Time To Analyze (MTTA) and Mean Time To Identify (MTTI) the affected assets. It is important to have metrics in place for each and any of the functions the SOC performs.

Well-defined processes and workflows for knowledge sharing and collaboration improve many if not all functions a SOC performs. Collaboration fosters knowledge sharing from itself, and will eventually benefit everyone in the team. In an operational setting, well-defined processes for event handling, case management and work shift changes will improve operations and will make it easier to repeat certain actions, eventually resulting in more automated processes, further increasing the SOC efficiency.

The security monitoring and incident response processes can be supported by technology. One has to think of communication platforms that foster collaboration and formal processes for handling events and incidents. Sometimes these capabilities are offered by a SIEM, but because these systems are not built specifically for this, the functionality may be lacking. It may then be necessary to invest in software that does provide these capabilities, which can come in many forms. An open source example is Fast Incident Response (FIR)[73], which has been developed by CERT Société Générale. Another example is Timesketch[74], which is a (proof of concept) tool for collaborative forensic timeline analysis.

Modern SOCs need a diverse set of experience and knowledge for operational success. Only performing compliance and security monitoring may be enough for certain organizations, but when more is at stake, more specialized skills are necessary. The diversity of the SOC operators skillsets can be increased by hiring from a more diverse group of people, including people with a background in security architecture engineering, programming, statistics, machine learning or reverse engineering, depending on the SOC functions. Educating existing staff is also a possibility, with the added benefit that this also keeps them fresh.

Your SOC staff can benefit greatly from having tools in place that support (automated) workflows and collaboration. Events and incidents that occur repeatedly can in some cases be coded in a playbook or even in software, allowing faster remediation. Increased collaboration, through real-time chats and ticketing systems for example, increases common understanding of the IT environment and can result in more efficient remediation. Demisto[75] is a vendor selling an example of software that increases collaboration and automation through a chatops approach: a central chat interface integrates with different types of security solutions, recording the actions performed by analysts and providing them with a single interface from which they can perform analyses. Phantom[76] also increases automation and performs security orchestration, through providing integrations with other security solutions and both proprietary

73 https://github.com/certsocietegenerale/FIR/
74 https://github.com/google/timesketch
75 https://www.demisto.com/
76 https://www.phantom.us/

and community playbooks for performing various security operations.

SOC INFRASTRUCTURE

The levels of data collection and/or fusion that can be realized with the currently deployed technologies is no longer sufficient to get a view of the entire IT environment. This means that data collection has to be tuned: prioritize data collection from sources related to the critical business processes and improve from there. The infrastructure should be able to at least aggregate log data. In order to get more visibility into the network it may be necessary to capture network traffic data (flows or full packet capture) and objects being transferred. Great visibility into endpoint systems across operating systems and devices can be realized with solutions like Tanium[77] and Lima Charlie[78].

The current technology stack does not give your security operators a complete view of the IT environment. Data collection and/or fusion capabilities of the currently deployed security solutions may not be sufficient. These can be improved by opting for modern solutions backed by scalable technologies. Examples of these include LogRhythm[79], Splunk[80] and the ELK stack[81], which all provide virtually limitless scaling with regards to storage and processing. These technologies often require specialist knowledge to deploy however.

Alternatively one can choose to use a Hadoop[82] distribution, such as those offered by MapR[83] and Cloudera[84], which also offer a lot of storage and processing scaling capabilities and information security use cases. These distributions often offer many integration and data access options, which allows for a great amount of flexibility. These include SQL interfaces to (non- or semi structured) data, built-in graph databases, machine learning engines and full text search. All of these can

77 https://www.tanium.com/
78 https://github.com/refractionPOINT/limacharlie
79 https://logrhythm.com/
80 http://www.splunk.com/
81 https://www.elastic.co/products
82 https://hadoop.apache.org/
83 https://www.mapr.com/
84 http://www.cloudera.com/content/www/en-us/solutions/information-security.html

then be combined with other software for example to visualize the available data and perform further analysis.

Not all of the functions performed by the SOC are properly being supported by the current infrastructure. It may be the case that the SIEM does not offer the right capabilities anymore and additional technology is necessary. When it is out of question that the current SIEM is going to be replaced or scaled up, the SOC can consider making use of other free, open source or paid extensions or alternatives to the SIEM. Some examples include the ELK stack[85] and MozDef[86]. Using additional tools to access the data already available in the SOC infrastructure may prove useful too. These can include custom integrations with the underlying databases, or plugging in Tableau[87] for instant access to diverse data sources.

Linked data analysis is an efficient method to detect adversaries and to investigate security incidents. The usage of linked data models, possibly supported by graph databases, makes efficient (visual) analysis of the data possible.

When operating any centralized data collection it is of utmost importance that access to that single repository is secure. Data collection should be performed securely, which means encryption should be deployed in-transit. When the data is at rest, it is possibly best to encrypt it also (encrypted at-rest), such that in case of loss of control over the data not all is lost. Additionally, authentication, authorization, accounting and audit of the data should be in place, so that the data that is stored in the system, can be accounted for. Performing security tests on the monitoring infrastructure is advised. When procuring new technology, determine the consequences the solution may have on the current infrastructure. VSAQ[88] is an interactive questionnaire application that can be used to assess the security programs of third parties.

85 https://www.elastic.co/products
86 https://github.com/jeffbryner/MozDef
87 https://www.tableau.com/
88 https://github.com/google/vsaq

ACRONYMS

API	Application Programming Interface
APT	Advanced Persistent Threat
AV	anti-virus
BA	Business Analytics
BI	Business Intelligence
CAPEX	Capital Expenditures
CEH	Certified Ethical Hacker
CEO	Chief Executive Officer
CERT	Computer Emergency Response Team
CFG	Confirmatory Focus Group
CIA	Confidentiality, Integrity and Availability
CISO	Chief Information Security Officer
CISSP	Certified Information Systems Security Professional
CND	Computer Network Defense
COA	course of action
CRISP-DM	CRoss-Industry Standard Process for Data Mining
CSV	comma-separated value
CVE	Common Vulnerabilities and Exposures
CybOX	Cyber Observable eXpression
DBIR	Data Breach Investigations Report
DDoS	Distributed Denial of Service
DLP	Data Loss Prevention
DS	Design Science
DsA	Descriptive Analytics
DSR	Design Science Research

DSRP	Design Science Research Process
EFG	Exploratory Focus Group
EMA	Enterprise Management Associates
ENISA	European Union Agency for Network and Information Security
ESG	Enterprise Strategy Group
FG	Focus Group
FIRST	Forum for Incident Response and Security Teams
FISMA	Federal Information Security Management Act
HP	Hewlett-Packard
HTML	HyperText Markup Language
HU	Hermeneutical Unit
ICS	Industrial Control System
ICT	Information and Communication Technology
IDC	International Data Corporation
IDS	Intrusion Detection System
IOC	Indicator of Compromise
IODEF	Incident Object Description Exchange Format
IP	Internet Protocol
IPS	Intrusion Prevention System
IR	Incident Response
IRT	Incident Response Team
IS	Information Systems
ISMS	Information Security Management System
ISP	Internet Service Provider
IoT	Internet-of-Things
IT	Information Technology
KPI	Key Performance Indicator
MBA	Master of Business Administration

MC	Mobile Computing
MSSP	Managed Security Services Provider
NCSC	Nationaal Cyber Security Centrum
NSM	Network Security Monitoring
NIST	National Institute of Standards and Technology
OCTAVE	Operationally Critical Threat, Asset, and Vulnerability Evaluation
OE	Operational Environment
OPEX	Operational Expenditures
OSINT	Open Source Intelligence
PCAP	packet capture
PD	Primary Document
PdA	Predictive Analytics
PsA	Prescriptive Analytics
SA	Security Analytics
SABSA	Sherwood Applied Business Security Architecture
SANS	SysAdmin, Audit, Network, and Security
SCADA	Supervisory Control and Data Acquisition
SIEM	Security Information and Event Management
SME	small and medium-sized enterprise
SOC	Security Operations Center
STIX	Structured Threat Information eXpression
SQL	Structured Query Language
SRA	Security Risk Assessment
SRM	Security Risk Management
TAL	Threat Agent Library
TAXII	Trusted Automated eXchange of Indicator Information
TI	Threat Intelligence
TTP	Tactics, Techniques and Procedures
VPN	Virtual Private Network

INDEX

BIBLIOGRAPHY

PEER-REVIEWED LITERATURE

[1] United States Public Law 107-347. *Federal Information Security Management Act (FISMA). Title III of the E-Government Act of 2002.* Dec. 2002. URL: http://www.gpo.gov/fdsys/pkg/PLAW-107publ347/html/PLAW-107publ347.htm (cit. on p. 11).

[2] Leman Akoglu, Hanghang Tong, and Danai Koutra. "Graph based anomaly detection and description: a survey". In: *Data Mining and Knowledge Discovery* 29.3 (2014), pp. 626–688. ISSN: 1573-756X. DOI: 10.1007/s10618-014-0365-y. URL: http://dx.doi.org/10.1007/s10618-014-0365-y (cit. on p. 53).

[3] Mamoun Alazab et al. "Global Security, Safety and Sustainability & e-Democracy: 7th International and 4th e-Democracy, Joint Conferences, ICGS3/e-Democracy 2011, Thessaloniki, Greece, August 24-26, 2011, Revised Selected Papers". In: ed. by Christos K. Georgiadis et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. Chap. Cybercrime: The Case of Obfuscated Malware, pp. 204–211. ISBN: 978-3-642-33448-1. DOI: 10.1007/978-3-642-33448-1_28. URL: http://dx.doi.org/10.1007/978-3-642-33448-1_28 (cit. on p. 39).

[4] Sean Barnum. *Standardizing cyber threat intelligence information with the Structured Threat Information eXpression (STIX™).* 2012 (cit. on p. 50).

[5] Sean Barnum et al. *The CybOX Language Specification*. https://cybox.mitre.org/language/specifications/CybOX_Language_Core_Specification_v1.0.pdf. Apr. 2012. URL: https://cybox.mitre.org/language/specifications/CybOX_Language_Core_Specification_v1.0.pdf (cit. on p. 50).

[6] Tim Bass. "Intrusion Detection Systems and Multisensor Data Fusion". In: *Commun. ACM* 43.4 (Apr. 2000), pp. 99–105. ISSN: 0001-0782. DOI: 10.1145/332051.332079. URL:

http://doi.acm.org/10.1145/332051.332079 (cit. on p. 51).

[7] S. Bhatt, P. K. Manadhata, and L. Zomlot. "The Operational Role of Security Information and Event Management Systems". In: *Security Privacy, IEEE* 12.5 (Sept. 2014), pp. 35–41. ISSN: 1540-7993. DOI: 10.1109/MSP.2014.103 (cit. on pp. 37, 38).

[8] The CEE Editorial Board. *Common Event Expression: Architecture Overview vo.5*. Tech. rep. May 2010. URL: https://cee.mitre.org/docs/CEE_Architecture_Overview-v0.5.pdf (cit. on p. 15).

[9] R. Buckminster Fuller. *World Design Science Decade 1965-1975, Phase I, Document 3, Comprehensive Thinking*. 1965 (cit. on p. 18).

[10] Sergio Caltagirone, Andrew Pendergast, and Christopher Betz. *The diamond model of intrusion analysis*. Tech. rep. DTIC Document, 2013. URL: https://www.threatconnect.com/wp-content/uploads/ThreatConnect-The-Diamond-Model-of-Intrusion-Analysis.pdf (cit. on p. 55).

[11] Richard Caralli et al. *Introducing OCTAVE Allegro: Improving the Information Security Risk Assessment Process*. Tech. rep. CMU/SEI-2007-TR-012. Pittsburgh, PA: Software Engineering Institute, Carnegie Mellon University, 2007. URL: http://resources.sei.cmu.edu/library/asset-view.cfm?AssetID=8419 (cit. on p. 11).

[12] Alvaro A. Cardenas, Saurabh Amin, and Shankar Sastry. "Secure Control: Towards Survivable Cyber-Physical Systems". In: *2013 IEEE 33rd International Conference on Distributed Computing Systems Workshops* (2008), pp. 495–500. ISSN: 1545-0678. DOI: http://doi.ieeecomputersociety.org/10.1109/ICDCS.Workshops.2008.40 (cit. on p. 37).

[13] Alvaro A. Cárdenas, Pratyusa K. Manadhata, and Sreeranga P. Rajan. "Big Data Analytics for Security Intelligence". In: *Cloud Security Alliance* September (2013), pp. 1–22. DOI: 10.1145/2666652.2666664. URL: https://downloads.cloudsecurityalliance.org/initiatives/bdwg/Big%5C_Data%5C_Analytics%5C_for%5C_Security%5C_Intelligence.pdf (cit. on pp. 2, 38).

[14] Alvaro A. Cárdenas, Pratyusa K. Manadhata, and Sreeranga P. Rajan. "Big Data Analytics for Security". In: *IEEE Security & Privacy* 11.6 (2013), pp. 74–76. ISSN: 1540-7993. DOI: http://doi.ieeecomputersociety.org/10.1109/MSP.2013.138 (cit. on p. 2).

[15] Timothy Casey, Patrick Koeberl, and Claire Vishik. "Defining Threat Agents: Towards a More Complete Threat Analysis". English. In: *ISSE 2010 Securing Electronic Business Processes*. Ed. by Norbert Pohlmann, Helmut Reimer, and Wolfgang Schneider. Vieweg+Teubner, 2011, pp. 214–225. ISBN: 978-3-8348-1438-8. DOI: 10.1007/978-3-8348-9788-6_21 (cit. on p. 9).

[16] James Cebula and Lisa Young. *A Taxonomy of Operational Cyber Security Risks*. Tech. rep. CMU/SEI-2010-TN-028. Pittsburgh, PA: Software Engineering Institute, Carnegie Mellon University, 2010. URL: http://resources.sei.cmu.edu/library/asset-view.cfm?AssetID=9395 (cit. on p. 10).

[17] Ping Chen, Lieven Desmet, and Christophe Huygens. "Communications and Multimedia Security: 15th IFIP TC 6/TC 11 International Conference, CMS 2014, Aveiro, Portugal, September 25-26, 2014. Proceedings". In: ed. by Bart Decker and André Zúquete. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014. Chap. A Study on Advanced Persistent Threats, pp. 63–72. ISBN: 978-3-662-44885-4. DOI: 10.1007/978-3-662-44885-4_5. URL: http://dx.doi.org/10.1007/978-3-662-44885-4_5 (cit. on pp. 39, 40).

[18] Yulia Cherdantseva and Jeremy Hilton. "A Reference Model of Information Assurance & Security". In: *Availability, Reliability and Security (ARES), 2013 Eighth International Conference on*. IEEE. 2013, pp. 546–555 (cit. on p. 8).

[19] Ramkumar Chinchani et al. "Towards a theory of insider threat assessment". In: *Dependable Systems and Networks, 2005. DSN 2005. Proceedings. International Conference on*. June 2005, pp. 108–117. DOI: 10.1109/DSN.2005.94 (cit. on p. 36).

[20] David Chismon and Martyn Ruks. *Threat Intelligence: Collecting, Analysing, Evaluating*. Tech. rep. MWR InfoSecurity, Mar. 2015. URL: https://www.cpni.gov.uk/Documents/

Publications/2015/23-March-2015-MWR_Threat_Intelligence_whitepaper-2015.pdf (cit. on pp. 49, 50).

[21] Anton Chuvakin, Kevin J. Schmidt, and Christopher Phillips. *Logging and Log Management: The Authoritative Guide to Understanding the Concepts Surrounding Logging and Log Management*. 1st Edition. Rockland, MA 02370, USA: Syngress, Dec. 2012, p. 460. ISBN: 1597496359,978-1597496353 (cit. on p. 15).

[22] Paul R. Cichonski et al. *NIST Special Publication 800-61 Rev 2: Computer Security Incident Handling Guide*. Tech. rep. Aug. 2012, p. 79. DOI: 10.6028/NIST.SP.800-61r2 (cit. on pp. 12, 53).

[23] Anne Cleven, Philipp Gubler, and Kai M. Hüner. "Design alternatives for the evaluation of design science research artifacts". In: *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology - DESRIST '09*. New York, New York, USA: ACM Press, 2009, p. 1. ISBN: 9781605584089. DOI: 10.1145/1555619.1555645. URL: http://portal.acm.org/citation.cfm?doid=1555619.1555645 (cit. on p. 23).

[24] Jonathan Clough. "Principles of cybercrime". In: 2nd edition. Cambridge, CB2 8BS, United Kingdom: Cambridge University Press, Sept. 2015. Chap. Cybercrime, p. 579. ISBN: 978-1107698161 (cit. on p. 39).

[25] Stephen Cobb. "Sizing Cybercrime: Incidents and Accidents, Hints and Allegiations". In: *Virus Bulletin Conference*. Sept. 2015, pp. 8–15. URL: https://www.virusbulletin.com/uploads/pdf/conference/vb2015/Cobb-VB2015.pdf (cit. on p. 39).

[26] Michael S. Collins. *Network Security Through Data Analysis: Building Situational Awareness*. 1st. Sebastopol, CA 95472, USA: O'Reilly Media, 2014. ISBN: 1449357903, 978-1449357900 (cit. on p. 15).

[27] Carl Colwill. "Human Factors in Information Security: The Insider Threat - Who Can You Trust These Days?" In: *Inf. Secur. Tech. Rep.* 14.4 (Nov. 2009), pp. 186–196. ISSN: 1363-4127. DOI: 10.1016/j.istr.2010.04.004. URL: http://dx.doi.org/10.1016/j.istr.2010.04.004 (cit. on p. 36).

[28] Julie Connolly et al. *The Trusted Automated eXchange of Indicator Information (TAXII™)*. https://taxii.mitre.org/about/documents/Introduction_to_TAXII_White_Paper_November_2012.pdf. Aug. 2012. URL: https://taxii.mitre.org/about/documents/Introduction_to_TAXII_White_Paper_November_2012.pdf (cit. on p. 50).

[29] John W. Creswell. "Research Design - Qualitative, Quantitative, and Mixed Methods Approaches - A Framework for Design". In: *Research design Qualitative quantitative and mixed methods approaches*. Thousand Oaks, California 91320: SAGE Publications, 2003. Chap. A Framewor, pp. 3–26. ISBN: 0761924418. DOI: 10.3109/08941939.2012.723954 (cit. on p. 18).

[30] John W. Creswell. *Research Design: Qualitative, Quantitative, and Mixed Approaches*. 3rd editio. Thousand Oaks, California 91320: SAGE Publications, Inc, 2009, p. 295. ISBN: 9781412965569. DOI: 10.1002/1521-3773(20010316)40:6<9823::AID-ANIE9823>3.3.CO;2-C (cit. on pp. 27, 28).

[31] R. Danyliw, J. Meijer, and Y. Demchenko. *The Incident Object Description Exchange Format*. RFC 5070. RFC Editor, Dec. 2007, pp. 1–92. URL: https://tools.ietf.org/html/rfc5070 (cit. on p. 50).

[32] Thomas H. Davenport and Jeanne G. Harris. *Competing on Analytics: The New Science of Winning*. 1st. Boston, MA, USA: Harvard Business School Press, 2007. ISBN: 1422103323, 9781422103326 (cit. on p. 16).

[33] Dursun Delen and Haluk Demirkan. "Data, information and analytics as services". In: *Decision Support Systems* 55.1 (2013), pp. 359–363. ISSN: 0167-9236. DOI: 10.1016/j.dss.2012.05.044. URL: http://www.sciencedirect.com/science/article/pii/S0167923612001558 (cit. on pp. 16, 17).

[34] Martin E. Dempsey. *Joint Publication 2-0: Joint Intelligence*. Tech. rep. Oct. 2013. URL: http://www.dtic.mil/doctrine/new_pubs/jp2_0.pdf (cit. on pp. 49, 58).

[35] D. E. Denning. "An Intrusion-Detection Model". In: *Software Engineering, IEEE Transactions on* SE-13.2 (Feb. 1987), pp. 222–232. ISSN: 0098-5589. DOI: 10.1109/TSE.1987.232894 (cit. on p. 1).

[36] Peter J. Denning. "A new social contract for research". In: *Communications of the ACM* 40.2 (Feb. 1997), pp. 132–134. ISSN: 00010782. DOI: 10.1145/253671.253755. URL: http://portal.acm.org/citation.cfm?doid=253671.253755 (cit. on p. 18).

[37] Zakir Durumeric, Eric Wustrow, and J. Alex Halderman. "ZMap: Fast Internet-Wide Scanning and its Security Applications". In: *Proceedings of the 22nd USENIX Security Symposium*. Aug. 2013 (cit. on p. 39).

[38] Zakir Durumeric et al. "A Search Engine Backed by Internet-Wide Scanning". In: *Proceedings of the 22Nd ACM SIGSAC Conference on Computer and Communications Security*. CCS '15. Denver, Colorado, USA: ACM, 2015, pp. 542–553. ISBN: 978-1-4503-3832-5. DOI: 10.1145/2810103.2813703. URL: http://doi.acm.org/10.1145/2810103.2813703 (cit. on p. 39).

[39] Prahlad Fogla and Wenke Lee. "Evading Network Anomaly Detection Systems: Formal Reasoning and Practical Techniques". In: *Proceedings of the 13th ACM Conference on Computer and Communications Security*. CCS '06. Alexandria, Virginia, USA: ACM, 2006, pp. 59–68. ISBN: 1-59593-518-5. DOI: 10.1145/1180405.1180414. URL: http://doi.acm.org/10.1145/1180405.1180414 (cit. on p. 51).

[40] Ulrik Franke and Joel Brynielsson. "Cyber situational awareness – A systematic review of the literature". English. In: *Computers & Security* 46.Complete (2014), pp. 18–31. DOI: 10.1016/j.cose.2014.06.008 (cit. on p. 57).

[41] P. García-Teodoro et al. "Anomaly-based network intrusion detection: Techniques, systems and challenges". In: *Computers & Security* 28.1–2 (2009), pp. 18–28. ISSN: 0167-4048. DOI: http://dx.doi.org/10.1016/j.cose.2008.08.003. URL: http://www.sciencedirect.com/science/article/pii/S0167404808000692 (cit. on p. 51).

[42] Marcus Gibson and David Arnott. "The Use of Focus Groups in Design Science Research". In: *Australasian Conference on Information Systems*. Toowoomba: Association for Information Systems, 2007, pp. 327–337. URL: http://aisel.aisnet.org/acis2007/14 (cit. on pp. 23, 25, 26, 28).

[43] Thomas L Greenbaum. *The handbook for focus group research*. 2nd edition. SAGE Publications, Inc, Nov. 1998, p. 280. ISBN: 978-0761912538 (cit. on p. 25).

[44] Dawn G. Gregg, Uday R. Kulkarni, and Ajay S. Vinzé. "Understanding the philosophical underpinnings of software engineering research in information systems". In: *Information Systems Frontiers* 3.2 (2001), pp. 169–183. ISSN: 1387-3326. URL: http://link.springer.com/article/10.1023/A:1011491322406 (cit. on p. 21).

[45] F.L. Greitzer et al. "Combating the Insider Cyber Threat". In: *Security & Privacy, IEEE* 6.1 (Jan. 2008), pp. 61–64. ISSN: 1540-7993. DOI: 10.1109/MSP.2008.8 (cit. on p. 36).

[46] Dina Hadžiosmanović et al. "Through the Eye of the PLC: Semantic Security Monitoring for Industrial Processes". In: *Proceedings of the 30th Annual Computer Security Applications Conference*. ACSAC '14. New Orleans, Louisiana, USA: ACM, 2014, pp. 126–135. ISBN: 978-1-4503-3005-3. DOI: 10.1145/2664243.2664277. URL: http://doi.acm.org/10.1145/2664243.2664277 (cit. on p. 37).

[47] Alan R. Hevner et al. "Design Science in Information Systems Research". In: *MIS Quarterly* 28.1 (2004), pp. 75–105. ISSN: 02767783. DOI: 10.2307/25148625. arXiv: /dl.acm.org/citation.cfm?id=2017212.2017217 [http:]. URL: http://dl.acm.org/citation.cfm?id=2017217 (cit. on pp. 5, 18, 19, 22, 24, 28, 32, 75).

[48] Eric M. Hutchins, Michael J. Cloppert, and Rohan M Amin. "Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains". In: *Leading Issues in Information Warfare & Security Research* 1 (2011), p. 80. URL: http://www.lockheedmartin.com/content/dam/lockheed/data/corporate/documents/LM-White-Paper-Intel-Driven-Defense.pdf (cit. on pp. 39, 55).

[49] Joint Task Force Transformation Initiative. *NIST Special Publication 800-39: Managing Information Security Risk: Organization, Mission, and Information System View*. Tech. rep. Gaithersburg, MD, United States, Mar. 2011 (cit. on p. 39).

[50] ISO/IEC. *Information technology – Security techniques – Evaluation criteria for IT security – Part 1: Introduction and general model.* ISO/IEC 15408-1:2009. Geneva, Switzerland: International Organization for Standardization, 2009. URL: http://www.iso.org/iso/catalogue_detail?csnumber=50341 (cit. on p. 11).

[51] ISO/IEC. *Information technology - Security techniques - Information security management systems - Overview and vocabulary.* ISO/IEC 27000:2009. Geneva, Switzerland: International Organization for Standardization, 2009. URL: http://www.iso.org/iso/catalogue_detail?csnumber=41933 (cit. on p. 7).

[52] ISO/IEC. *Information technology – Security techniques – Code of practice for information security management*. ISO/IEC 27002:2013. Geneva, Switzerland: International Organization for Standardization, 2013. URL: http://www.iso.org/iso/catalogue_detail?csnumber=54533 (cit. on p. 8).

[53] ISO/IEC. *Information technology - Security techniques - Information security management systems - Requirements.* ISO/IEC 27000:2013. Geneva, Switzerland: International Organization for Standardization, 2013. URL: http://www.iso.org/iso/catalogue_detail?csnumber=41933 (cit. on p. 8).

[54] Jay Jacobs and Bob Rudis. *Data-Driven Security: Analysis, Visualization and Dashboards*. 1st edition. Hoboken, NJ, USA: Wiley, Apr. 2014, p. 352. ISBN: 978-1-118-79372-5 (cit. on p. 54).

[55] Sushil Jajodia et al. *Cyber situational awareness: Issues and research*. Ed. by Sushil (George Mason University) Jajodia et al. Vol. 46. New York: Springer, 2010, p. 249. ISBN: 9781441901392. DOI: 10.1007/978-1-4419-0140-8 (cit. on p. 48).

[56] Andrew Jaquith. *Security Metrics: Replacing Fear, Uncertainty, and Doubt*. 1st edition. Indianapolis, Indiana 46240, US: Addison-Wesley Professional, Apr. 2007, p. 336. ISBN: 978-0321349989 (cit. on pp. 49, 56).

[57] Paul Johannesson and Erik Perjons. *An Introduction to Design Science*. Cham: Springer International Publishing, 2014, p. 197. ISBN: 978-3-319-10631-1. DOI: 10.1007/978-

3-319-10632-8. URL: http://link.springer.com/10.1007/978-3-319-10632-8 (cit. on pp. 23, 27, 28).

[58] V. Jyothsna, V. V. Rama Prasad, and K. Munivara Prasad. "A review of anomaly based intrusion detection systems". In: *International Journal of Computer Applications* 28.7 (2011), pp. 26–35 (cit. on p. 51).

[59] Richard L. Kissel. *Glossary of Key Information Security Terms*. Tech. rep. June 2013. URL: http://www.nist.gov/manuscript-publication-search.cfm?pub_id=913810 (cit. on p. 53).

[60] Gabriel Klein, Heiko Günther, and Susan Träber. "Future Security: 7th Security Research Conference, Future Security 2012, Bonn, Germany, September 4-6, 2012. Proceedings". In: ed. by Nils Aschenbruck et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. Chap. Modularizing Cyber Defense Situational Awareness – Technical Integration before Human Understanding, pp. 307–310. ISBN: 978-3-642-33161-9. DOI: 10.1007/978-3-642-33161-9_46. URL: http://dx.doi.org/10.1007/978-3-642-33161-9_46 (cit. on p. 49).

[61] I. Kotenko and E. Novikova. "Visualization of Security Metrics for Cyber Situation Awareness". In: *Availability, Reliability and Security (ARES), 2014 Ninth International Conference on*. Sept. 2014, pp. 506–513. DOI: 10.1109/ARES.2014.75 (cit. on p. 49).

[62] Vadim Kotov and Fabio Massacci. "Engineering Secure Software and Systems: 5th International Symposium, ESSoS 2013, Paris, France, February 27 - March 1, 2013. Proceedings". In: ed. by Jan Jürjens, Benjamin Livshits, and Riccardo Scandariato. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. Chap. Anatomy of Exploit Kits, pp. 181–196. ISBN: 978-3-642-36563-8. DOI: 10.1007/978-3-642-36563-8_13. URL: http://dx.doi.org/10.1007/978-3-642-36563-8_13 (cit. on p. 38).

[63] Douglas J Landoll. *The Security Risk Assessment Handbook: A Complete Guide for Performing Security Risk Assessments*. Auerbach Publications, 2005, pp. xxi + 473. URL: http://www.amazon.com/Security-Risk-Assessment-Handbook-Assessments/dp/0849329981 (cit. on p. 10).

[64] Hung-Jen Liao et al. "Intrusion detection system: A comprehensive review". In: *Journal of Network and Computer Applications* 36.1 (2013), pp. 16–24. ISSN: 1084-8045. DOI: http://dx.doi.org/10.1016/j.jnca.2012.09.004. URL: http://www.sciencedirect.com/science/article/pii/S1084804512001944 (cit. on p. 51).

[65] Irv Lustig et al. "The Analytics Journey". In: *Analytics Magazine* (Nov. 2010), pp. 11–18. URL: http://www.analytics-magazine.org/november-december-2010/54-the-analytics-journey (cit. on pp. 16, 17).

[66] Derek Manky. "Cybercrime as a service: a very modern business". In: *Computer Fraud & Security* 2013.6 (2013), pp. 9–13. ISSN: 1361-3723. DOI: http://dx.doi.org/10.1016/S1361-3723(13)70053-8. URL: http://www.sciencedirect.com/science/article/pii/S1361372313700538 (cit. on p. 39).

[67] Raffael Marty. *Applied Security Visualization*. 1st ed. Indianapolis, Indiana 46240, US: Addison-Wesley Professional, 2008. ISBN: 0321510100, 9780321510105 (cit. on p. 49).

[68] Raffael Marty. *The Security Data Lake: Leveraging Big Data Technologies to Build a Common Data Repository for Security*. Ed. by Laurel Ruma and Shannon Cutt. 1st edition. 1005 Gravenstein Highway North, Sebastopol, CA 95472: O'Reilly Media, Inc., 2015, p. 36. ISBN: 9781491927731 (cit. on p. 84).

[69] James P McDermott. "Attack net penetration testing". In: *Proceedings of the 2000 workshop on New security paradigms*. ACM. 2001, pp. 15–21 (cit. on p. 9).

[70] David Miller et al. *Security information and event management (SIEM) implementation*. 1st edition. Network Pro Library. McGraw-Hill Education, Oct. 2010. ISBN: 978-0071701099 (cit. on p. 2).

[71] David L. Morgan. *Focus Groups as Qualitative Research*. 1st edition. Vol. 16. Qualitative Research Methods Series. Thousand Oaks, California 91320: SAGE Publications, Inc., 1988. ISBN: 0-8039-3209-1 (cit. on pp. 23, 25, 27, 28).

[72] David. L. Morgan. "Planning and Research Design for Focus Groups". In: *Focus Groups as Qualitative Research*. 2455 Teller Road, Thousand Oaks California 91320 United States of America: SAGE Publications, Inc., 1997. Chap. 4, pp. 32–45. ISBN: 9780761903437. DOI: 10.4135/9781412984287.n4. URL: http://srmo.sagepub.com/view/focus-groups-as-qualitative-research/n4.xml (cit. on p. 25).

[73] M. G. Mullen. *Joint Publication 3-0: Joint Operations*. Tech. rep. Aug. 2011. URL: http://www.dtic.mil/doctrine/new_pubs/jp3_0.pdf (cit. on p. 49).

[74] Committee on National Security Systems. *National Information Assurance (IA) Glossary*. Fort George G. Meade, MD: Committee on National Security Systems, Apr. 2010 (cit. on p. 10).

[75] Evgenia Novikova and Igor Kotenko. "Analytical visualization techniques for security information and event management". In: *Proceedings of the 2013 21st Euromicro International Conference on Parallel, Distributed, and Network-Based Processing, PDP 2013* (2013), pp. 519–525. DOI: 10.1109/PDP.2013.84 (cit. on p. 49).

[76] Kathryn Parsons et al. *Human Factors and Information Security: Individual, Culture and Security Environment*. Tech. rep. Command, Control, Communications, Intelligence Division DSTO Defence Science, and Technology Organisation, PO Box 1500 Edinburgh, South Australia 5111, Australia, Oct. 2010. URL: http://www.dtic.mil/dtic/tr/fulltext/u2/a535944.pdf (cit. on p. 36).

[77] Ken Peffers et al. "The Design Science Research Process: A Model for Producing and Presenting Information Systems Research". In: *The Proceedings of Design Research in Information Systems and Technology DESRIST'06* 24 (2006), pp. 83–106. URL: http://geni15.wrsc.org/sites/default/files/documents/000designscresearchproc_desrist_2006.pdf (cit. on pp. 4, 5, 18, 20, 29, 30).

[78] Jan Pries-Heje, Richard Baskerville, and John Venable. "Strategies for design science research evaluation". In: *Proceedings of the 16th European Conference on Information Systems*. Vol. 16. 2004. Galway, Ireland, 2008, pp. 255–266. ISBN: 9780955315923. URL: http://aisel.aisnet.org/ecis2008/87/ (cit. on p. 23).

[79] Chris Sanders and Jason Smith. *Applied Network Security Monitoring: Collection, Detection, and Analysis*. 1st Edition. Syngress, Dec. 2013, p. 496. ISBN: 0124172083,978-0124172081 (cit. on p. 15).

[80] José Jair Santanna et al. "Booters—An analysis of DDoS-as-a-service attacks". In: *Integrated Network Management (IM), 2015 IFIP/IEEE International Symposium on*. IEEE. 2015, pp. 243–251 (cit. on p. 39).

[81] Bruce Schneier. *Secrets & Lies: Digital Security in a Networked World*. 1st. New York, NY, USA: John Wiley & Sons, Inc., 2000. ISBN: 0471253111 (cit. on p. 56).

[82] Lui Sha et al. "Cyber-Physical Systems: A New Frontier". In: *Sensor Networks, Ubiquitous and Trustworthy Computing, 2008. SUTC '08. IEEE International Conference on*. June 2008, pp. 1–9. DOI: 10.1109/SUTC.2008.85 (cit. on p. 37).

[83] Colin Shearer. "The CRISP-DM model: the new blueprint for data mining". In: *Journal of Data Warehousing* 5.4 (2000), pp. 13–22 (cit. on p. 52).

[84] Mario Silic and Andrea Back. "Shadow IT - A View from Behind the Curtain". In: *Comput. Secur.* 45 (Sept. 2014), pp. 274–283. ISSN: 0167-4048. DOI: 10.1016/j.cose.2014.06.007. URL: http://dx.doi.org/10.1016/j.cose.2014.06.007 (cit. on p. 36).

[85] Herbert A. Simon. *The Sciences of the Artificial (3rd Ed.)* Cambridge, MA, USA: MIT Press, Sept. 1996. ISBN: 0-262-69191-4 (cit. on p. 18).

[86] Robin Sommer and Vern Paxson. "Outside the Closed World: On Using Machine Learning for Network Intrusion Detection". In: *Security and Privacy (SP), 2010 IEEE Symposium on*. May 2010, pp. 305–316. DOI: 10.1109/SP.2010.25 (cit. on pp. 51, 53, 55).

[87] National Institute of Standards and Technology (NIST). *NIST Special Publication 800-53 Revision 4: Security and Privacy Controls for Federal Information Systems and Organizations*. Gaithersburg, MD: National Institute of Standards and Technology, Apr. 2013. DOI: 10.6028/NIST.SP.800-53r4 (cit. on pp. 8, 9, 11).

[88] Prem N. Stewart David W. Shamdasani. *Focus Groups: Theory and Practice*. 3rd edition. Applied Social Research Methods. Thousand Oaks, California 91320: SAGE Publications, Inc, 2015, p. 244. ISBN: 9781452270982 (cit. on p. 25).

[89] Gary Stoneburner, Alice Y. Goguen, and Alexis Feringa. *SP 800-30. Risk Management Guide for Information Technology Systems*. Tech. rep. Gaithersburg, MD, United States, 2002 (cit. on p. 10).

[90] Gary Stoneburner, Clark Hayden, and Alexis Feringa. *SP 800-27 Rev. A. Engineering Principles for Information Technology Security (A Baseline for Achieving Security), Revision A*. Tech. rep. Gaithersburg, MD, United States, 2004 (cit. on p. 9).

[91] W. W. Streilein et al. "Cyber situational awareness through operational streaming analysis". In: *MILITARY COMMUNICATIONS CONFERENCE, 2011 - MILCOM 2011*. Nov. 2011, pp. 1152–1157. DOI: 10.1109/MILCOM.2011.6127455 (cit. on p. 49).

[92] Frank Swiderski and Window Snyder. *Threat Modeling*. Redmond, WA, USA: Microsoft Press, 2004. ISBN: 0735619913 (cit. on p. 9).

[93] Mark Talabis et al. *Information Security Analytics: Finding Security Insights, Patterns, and Anomalies in Big Data*. 1st edition. Rockland, MA 02370, USA: Syngress, Nov. 2014, p. 171. ISBN: 9780128002070 (cit. on p. 54).

[94] Herbert H. Thompson. "Application Penetration Testing". In: *IEEE Security and Privacy* 3.1 (Jan. 2005), pp. 66–69. ISSN: 1540-7993. DOI: 10.1109/MSP.2005.3. URL: http://dx.doi.org/10.1109/MSP.2005.3 (cit. on p. 9).

[95] Monica Chiarini Tremblay, Alan R. Hevner, and Donald J. Berndt. "Focus Groups for Artifact Refinement and Evaluation in Design Research". In: *Design Research in Information Systems*. Vol. 22. Integrated Series in Information Systems 1. Springer US, 2010, pp. 121–143. ISBN: 978-1-4419-5652-1. DOI: 10.1007/978-1-4419-5653-8_10. URL: http://link.springer.com/10.1007/978-1-4419-5653-8%5C_10 (cit. on pp. 25, 27).

[96] Chih-Fong Tsai et al. "Intrusion detection by machine learning: A review". In: *Expert Systems with Applications* 36.10 (2009), pp. 11994–12000. ISSN: 0957-4174. DOI: http://dx.doi.org/10.1016/j.eswa.2009.05.029. URL: http://www.sciencedirect.com/science/article/pii/S0957417409004801 (cit. on p. 51).

[97] Vijay K. Vaishnavi and William Kuechler. *Design Science Research Methods and Patterns: Innovating Information and Communication Technology*. 1st edition. Boston, MA, USA: Auerbach Publications, Oct. 2007. ISBN: 978-1-4200-5932-8. DOI: 10.1201/9781420059335. URL: http://www.crcnetbase.com/doi/book/10.1201/9781420059335 (cit. on pp. 20, 21).

[98] Vijay Vaishnavi and Bill Kuechler. "Design Science Research in Information Systems". In: *Association for Information Systems* (2004), p. 45. URL: http://desrist.org/desrist/content/design-science-research-in-information-systems.pdf (cit. on p. 20).

[99] John Venable, Jan Pries-Heje, and Richard Baskerville. "A Comprehensive Framework for Evaluation in Design Science Research". In: *Design Science Research in Information Systems. Advances in Theory and Practice* (2012), pp. 423–438. ISSN: 0960-085X. DOI: 10.1007/978-3-642-29863-9_31 (cit. on pp. 23, 32).

[100] Michael Vidulich et al. *Situation Awareness: Papers and Annotated Biliography*. Tech. rep. DTIC Document, June 1994, p. 174. URL: http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA284752 (cit. on p. 48).

[101] Nikos Virvilis and Dimitris Gritzalis. "The Big Four - What We Did Wrong in Advanced Persistent Threat Detection?" In: *Proceedings of the 2013 International Conference on Availability, Reliability and Security*. ARES '13. Washington, DC, USA: IEEE Computer Society, 2013, pp. 248–254. ISBN: 978-0-7695-5008-4. DOI: 10.1109/ARES.2013.32. URL: http://dx.doi.org/10.1109/ARES.2013.32 (cit. on p. 52).

[102] Ron Weber. "Editor's Comment: Still Desperately Seeking the IT Artifact". In: *MIS Quarterly* 27.2 (June 2003), pp. iii–xi. ISSN: 0276-7783. URL: http://dl.acm.org/citation.cfm?id=2017189.2017190 (cit. on p. 18).

[103] Moira J. West-Brown et al. *Handbook for Computer Security Incident Response Teams (CSIRTS)*. 2nd Edition. Carnegie Mellon University, Apr. 2003. URL: https://www.sei.cmu.edu/reports/03hb002.pdf (cit. on pp. 12, 13).

[104] Roel J. Wieringa. *Design Science Methodology for Information Systems and Software Engineering*. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg, 2014, p. 493. ISBN: 978-3-662-43838-1. DOI: 10.1007/978-3-662-43839-8. URL: http://portal.acm.org/citation.cfm?doid=1810295.1810446 (cit. on p. 32).

[105] Aaron Woody. *Enterprise Security: A Data-Centric Approach to Securing the Enterprise*. 1st edition. Packt Publishing, Feb. 2013, p. 324. ISBN: 978-1849685962 (cit. on p. 53).

[106] Carson Zimmerman. *Ten Strategies of a World-Class Cybersecurity Operations Center*. 1st edition. 202 Burlington Road, Bedford, MA 01730-1420 (781) 271-2000: The MITRE Corporation, 2014. ISBN: 978-0-692-24310-7 (cit. on pp. 2, 11–13, 37, 54, 55).

[107] Richard Zuech, Taghi M. Khoshgoftaar, and Randall Wald. "Intrusion detection and Big Heterogeneous Data: a Survey". English. In: *Journal of Big Data* 2.1, 3 (2015). DOI: 10.1186/s40537-015-0013-4 (cit. on pp. 51, 57).

OTHER LITERATURE

[108] Lillian Ablon, Martin C. Libicki, and Andrea A. Golay. *Markets for Cybercrime Tools and Stolen Data: Hackers' Bazaar*. Tech. rep. Santa Monica, CA: RAND Corporation, July 2014. URL: http://www.rand.org/pubs/research_reports/RR610.html (cit. on p. 39).

[109] Ann Bednarz. *Shortage of security pros worsens*. http://www.networkworld.com/article/2893365/security0/shortage-of-security-pros-worsens.html. [Online; accessed April 5th, 2016]. Mar. 2015 (cit. on p. 54).

[110] Timothy Casey. *Threat Agent Library Helps Identify Information Security Risks*. Tech. rep. Intel Corporation, Sept. 2007. URL: https://communities.intel.com/docs/DOC-1151 (cit. on p. 9).

[111] Anton Chuvakin. *SIEM architecture and operational processes [Webinar]*. http://searchsecurity.techtarget.com/video/SIEM-architecture-and-operational-processes. June 2015. URL: http://searchsecurity.techtarget.com/video/SIEM-architecture-and-operational-processes (cit. on pp. 2, 38).

[112] Anton Chuvakin. *Starting A SIEM Project from Vendor Use Case Content: WIN or FAIL?* Dec. 2015. URL: http://blogs.gartner.com/anton-chuvakin/2015/12/02/starting-a-siem-project-from-vendor-use-case-content-win-or-fail/ (cit. on p. 38).

[113] CrowdStrike. *2015 Global Threat Report*. Tech. rep. CrowdStrike, Feb. 2016. URL: http://go.crowdstrike.com/rs/281-OBQ-266/images/15GlobalThreatReport.pdf (cit. on p. 37).

[114] Dell. *Dell Security Annual Threat Report*. Tech. rep. Dell Security, Apr. 2015. URL: http://www.dell.com/learn/us/en/uscorp1/press-releases/2015-04-13-dell-annual-threat-report (cit. on p. 1).

[115] John Friedman and Mark Bouchard. *Definitive Guide to Cyber Threat Intelligence: Using Knowledge about Adversaries to Win the War against Targeted Attacks*. 1997 Annapolis Exchange Parkway, Suite 300, Annapolis, MD 21401: CyberEdge Group, LLC, Sept. 2015. ISBN: 978-0-9961827-1-3 (cit. on pp. 49, 50).

[116] John Gantz and David Reinsel. *Extracting Value from Chaos*. Tech. rep. 5 Speen Street Framingham, MA 01701 USA: International Data Corporation (IDC), 2011. URL: http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf (cit. on p. 3).

[117] Art Gilliland. "Stop Looking for the Silver Bullet: Start Thinking Like a Bad Guy". RSA Conference USA. Feb. 2014. URL: http://www.rsaconference.com/events/us14/agenda/sessions/1344/stop-looking-for-the-silver-bullet-start-thinking (cit. on p. 1).

[118] Ely Kahn. *Building a Next Gen SOC [Webinar]*. http://info.sqrrl.com/sqrrl-october-webinar-next-generation-soc. Oct. 2015. URL: http://info.sqrrl.com/sqrrl-october-webinar-next-generation-soc (cit. on p. 14).

[119] Kelly M. Kavanagh and Oliver Rochford. *Magic Quadrant for Security Information and Event Management*. Tech. rep. Stamford CT, U.S.A.: Gartner Research, July 2015 (cit. on p. 2).

[120] Robert Lemos. *The Hunt for Data Analytics: Is Your SIEM on the Endangered List?* http://searchsecurity.techtarget.com/feature/The-hunt-for-data-analytics-Is-your-SIEM-on-the-endangered-list. 275 Grove Street, Newton, MA 02466, Mar. 2015 (cit. on p. 36).

[121] Mandiant. *Are you ready to respond? Evaluate and Improve Your Ability to Respond to the Next Attack*. Tech. rep. Mandiant (FireEye, Inc.), 2015. URL: https://dl.mandiant.com/EE/library/WP_Are_You_Ready_To_Respond.pdf (cit. on p. 56).

[122] Mandiant. *M-Trends 2015: A View From The Front Lines*. Tech. rep. Mandiant, Feb. 2015. URL: http://www2.fireeye.com/rs/fireye/images/rpt-m-trends-2015.pdf (cit. on p. 1).

[123] Rob McMillan. *Definition: Threat Intelligence*. Tech. rep. Gartner, May 2013. URL: https://www.gartner.com/doc/2487216? (cit. on p. 49).

[124] David Monahan. *The Evolution of Data Driven Security*. Tech. rep. Enterprise Management Associates, June 2014, pp. 1–30. URL: http://research.enterprisemanagement.com/rs/ema/images/EMA_EDDS_2014_RR.pdf (cit. on pp. 36–38).

[125] David Monahan. *Data-Driven Security Reloaded*. Tech. rep. Enterprise Management Associates, Apr. 2015, pp. 1–8. URL: https://cdn2.hubspot.net/hubfs/208516/Assets/EMA-Prelert-DataDrivenSecurityReloaded-Summary_Apr2015.pdf (cit. on p. 37).

[126] Linda Musthaler. *Security analytics will be the next big thing in IT security*. http://www.networkworld.com/article/2166806/infrastructure-management/security-analytics-will-be-the-next-big-thing-in-it-security.html. [Online; accessed March 18, 2015]. May 2013 (cit. on p. 2).

[127] European Union Agency for Network and Information Security (ENISA). *ENISA Threat Landscape 2014 - Overview of current and emerging cyber-threats*. https://www.enisa.europa.eu/activities/risk-management/evolving-threat-environment/enisa-threat-landscape/enisa-threat-landscape-2014. Jan. 2015 (cit. on p. 9).

[128] European Union Agency for Network and Information Security (ENISA). *ENISA Threat Landscape 2015*. https://www.enisa.europa.eu/activities/risk-management/evolving-threat-environment/enisa-threat-landscape/etl2015. Jan. 2016 (cit. on p. 37).

[129] Jon Oltsik. *The Big Data Security Analytics Era Is Here*. Tech. rep. Enterprise Security Group, Jan. 2013, pp. 1–13. URL: https://www.emc.com/collateral/analyst-reports/security-analytics-esg-ar.pdf (cit. on pp. 36, 38).

[130] Jon Oltsik. *The Evolution of Big Data Security Analytics Technology*. Tech. rep. Enterprise Security Group, Mar. 2013, pp. 1–18. URL: http://www.esgnext.com/c/701C0000000eY2H/pdf/ESG_MLR_Big_Data_Security_Analytics_Mar_2013.pdf (cit. on pp. 38, 51, 57).

[131] Jon Oltsik. *Enterprise Organizations Need Contextual Security Analytics*. Tech. rep. Enterprise Security Group, Oct. 2014, pp. 1–5. URL: https://www.lancope.com/sites/default/files/ESG-Brief-Lancope-FINAL_0.pdf (cit. on pp. 37, 38).

[132] Jon Oltsik. *An Analytics-based Approach to Cybersecurity*. Tech. rep. Enterprise Security Group, May 2015, pp. 1–6. URL: https://www.hexiscyber.com/sites/default/

files/An%20Analytics-based%20Approach%20to%20Cybersecurity-2015_0.pdf (cit. on p. 38).

[133] Sreeranga Rajan, Wilco van Ginkel, and Neel Sundaresan. *Top Ten Big Data Security And Privacy Challenges*. Tech. rep. Cloud Security Alliance, Nov. 2012. URL: http://www.isaca.org/groups/professional-english/big-data/groupdocuments/big_data_top_ten_v1.pdf (cit. on p. 57).

[134] Global Research and Analysis Team (GReAT). *Energetic Bear — Crouching Yeti*. Tech. rep. Kaspersky Lab, July 2014. URL: https://cdn.securelist.com/files/2014/07/EB-YetiJuly2014-Public.pdf (cit. on p. 40).

[135] Global Research and Analysis Team (GReAT). *Equation Group - Questions and Answers*. Tech. rep. Kaspersky Lab, Feb. 2015. URL: https://securelist.com/files/2015/02/Equation_group_questions_and_answers.pdf (cit. on p. 1).

[136] Mike Rothman and Adrian Lane. *Security Management 2.5: Replacing Your SIEM Yet?* Tech. rep. 515 E. Carefree Highway Suite #766, Phoenix, AZ 85085: Securosis, Feb. 2014. URL: https://securosis.com/assets/library/reports/SecurityManagement2.5_FINAL-multi.pdf (cit. on pp. 36–38).

[137] Dave Shackleford. *Analytics and Intelligence Survey 2014*. Tech. rep. SANS Institute, Oct. 2014. URL: http://www.sans.org/reading-room/whitepapers/analyst/analytics-intelligence-survey-2014-35507 (cit. on pp. 1, 36, 38).

[138] Dave Shackleford. *Analytics and Intelligence Survey 2015*. Tech. rep. SANS Institute, Nov. 2015. URL: https://www.sans.org/reading-room/whitepapers/analyst/2015-analytics-intelligence-survey-36432 (cit. on pp. 36, 38).

[139] Alissa Torres. *Building a World-Class Security Operations Center: A Roadmap*. Tech. rep. SANS Institute, May 2015. URL: https://www.sans.org/reading-room/whitepapers/analyst/building-world-class-security-operations-center-roadmap-35907 (cit. on p. 13).

[140] Verizon. *2014 Data Breach Investigations Report*. Tech. rep. Verizon, Apr. 2014. URL: http://verizonenterprise.com (cit. on p. 1).

[141] Verizon. *2015 Data Breach Investigations Report*. Tech. rep. Verizon, Apr. 2015. URL: http://verizonenterprise.com (cit. on p. 1).

DECLARATION

I hereby declare that this thesis is entirely my own work and that any additional sources of information have been duly cited.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this thesis has not been submitted for a higher degree to any University or Institution other than the University of Twente.

Enschede, 6 June 2016

Herman Slatman