NTNU
Norwegian University of
Science and Technology

Improving Security Posture by Learning from Intrusions

Anders Nese

Master of Science in Communication Technology
Submission date: June 2018
Supervisor: Karin Bernsmed, IIK

Norwegian University of Science and Technology
Department of Information Security and Communication Technology

Title: Improving Security Posture by Learning from Intrusions

Student: Anders Nese

Problem description:

Several research papers reveal that organisations lack structured approaches for learning from incidents [56, 35]. Although learning from incidents is emphasised in well-adopted standards and guidelines on incident response [34, 16, 58], research reveals that organisations find it difficult to implement this concept in practice [73, 29]. This results in organisations missing out on great opportunities to improve their security posture. In this thesis, we will look at a subtype of incidents called intrusions, and how organisations can improve their security posture by proactively handle and investigate all intrusions, including failed intrusion attempts. We denote intrusions, as in the technical report of the *Diamond Model of Intrusion Analysis* [14], to be '*all malicious and nefarious activity targeting computer systems and networks*'.

The research question of the thesis will be as follows:

> How can organisations leverage lessons learned about historical intrusions to improve their security posture?

The main research question is further divided into sub-questions. A solid basis for discussing the main research question will be established by answering the following questions:

- How can intrusion analysis help expedite prevention and detection of intrusions?
- How can indicators be used to identify previously undetected intrusions?

In order to answer the research questions, information about how organisations leverage lessons learned after intrusions will be gathered. Experiences from a variety of intrusions will be systematised through a case study, semi-structured interviews with industry experts will be conducted, and a study of how organisations can better utilise lessons learned after intrusions will be performed.

Responsible professor: Karin Bernsmed, IIK

Supervisor: Maria Kjærland-Haga, External

Abstract

Previous research have found that organisations lack structured approaches for learning from incidents, which results in organisations missing out on opportunities to improve their security posture. In this thesis, qualitative interviews with industry experts are used in combination with a case study to explore how organisations could leverage intrusions to improve their security posture. Findings from the interviews indicate that there is a lack of structured methods for organisations to learn from intrusions integrating double-loop learning, proactive discovery and information sharing. There are, nonetheless, models that structure *either* organisational learning *or* intelligence-driven active defence. One consequence is that high-value intelligence generated from intrusion data is not used effectively, or not used at all, when generating threat hunting hypothesises. Further, without a structured approach for sharing information, stakeholders that could have acted on that intelligence are instead making less informed decisions.

To overcome these shortcomings, we introduce a model integrating post-incident activities with intelligence, adversary discovery and information sharing. The purpose of this model is to explicate how data, information and knowledge from intrusions could be used in a structured approach for proactive defensive operations and improved information flows. We argue that widening the scope of incident response standards and guidelines to embrace proactive defence principles, such as learning from intrusions, intelligence and adversary discovery, would aid organisations in structuring their holistic approach to cyber security and make it easier for them to adopt an active defence approach.

Sammendrag

Tidligere forskning viser at organisasjoner mangler strukturerte tilnærminger for å lære av hendelser, noe som resulterer i at organisasjoner går glipp av muligheter til å forbedre sin digitale sikkerhet. Dette er en kvalitativ undersøkelse, der det er gjennomført åtte semistrukturerte intervju med industrieksperter på hendelseshåndtering. De kvalitative intervjuene blir brukt i kombinasjon med resultater fra en case studie for å utforske hvordan organisasjoner kan lære av digitale angrep. Undersøkelsen indikerer at det mangler strukturerte metoder som integrerer dyp læring, proaktiv deteksjon og deling av informasjon. Slike metoder kunne vært brukt av organisasjoner for å bedre lære av digitale angrep. Det er likevel modeller som strukturerer enten organisatorisk læring eller etterretningsdrevet proaktivt forsvar. En konsekvens av dette er at kvalitetsettereretning generert av data fra digitale angrep ikke blir brukt effektivt, eller ikke brukt i det hele tatt, når det arbeides med proaktiv trusseldeteksjon. Uten en strukturert tilnærming til deling kan interessenter som kunne ha handlet på denne etterretningen i stedet ende opp med å ta uinformerte beslutninger.

For å løse disse problemene introduserer vi en modell som integrerer hendelseshåndtering med etterretning, proaktiv trusseldeteksjon og informasjonsdeling. Formålet med denne modellen er å forklare hvordan data, informasjon og kunnskap fra digitale angrep kan brukes i en strukturert tilnærming til proaktivt forsvarsarbeid og bedre informasjonsflyten i bedrifter. Vi mener at en utvidelse av omfanget i standarder og retningslinjer for hendelseshåndtering til å omfavne proaktive forsvarsprinsipper, som læring fra digitale angrep, etterretning og proaktiv trusseldeteksjon, vil hjelpe organisasjoner i å strukturere deres holistiske tilnærming til digital sikkerhet og gjøre det enklere for dem å ivareta proaktive forsvarsmekanismer.

Preface

This master's thesis is submitted to the Norwegian University of Science and Technology (NTNU) as the final part of the five-year Master of Science in Communication Technology program at the Department of Information Security and Communication Technology (IIK).

I would like to thank my supervisor Maria Kjærland-Haga and professor Karin Bernsmed for valuable comments and guidance throughout this work. Their contributions have been invaluable. I would also like to thank all participants who took the time to participate in the interviews.

Trondheim, June, 2018
Anders Nese

Contents

List of Figures

List of Acronyms

ACDC Active Cyber Defense Cycle.

APT Advanced Persistent Threat.

AV Antivirus.

C2 Command and Control.

CERT Computer Emergency Response Team.

CSIRT Computer Security Incident Response Team.

CTI Cyber Threat Intelligence.

DMZ Demilitarized Zone.

DNS Domain Name System.

F3EAD Find, Fix, Finish, Exploit, Analyse and Disseminate.

HMI Human Machine Interface.

ICS Industrial Control Systems.

IDS Intrusion Detection System.

IOC Indicator of Compromise.

IPIRO Intel-Pervaded Incident Response Operations.

IPS Intrusion Prevention System.

IRT Incident Response Team.

LFI Learning from Incidents.

NIST National Institute of Standards and Technology.

RDP Remote Desktop Protocol.

ROI Return on Investment.

SCADA Supervisory Control And Data Acquisition.

SIS Safety Instrumented System.

SOC Security Operations Center.

STIX Structured Threat Information eXpression.

TAXII Trusted Automated eXchange of Indicator Information.

TLP Traffic Light Protocol.

TTPs Tactics, Techniques and Procedures.

VERIS Vocabulary for Event Recording and Incident Sharing.

Chapter 1
Introduction

[...] there are two kinds of big companies [...] there are those who've been hacked [...] and those who don't know they've been hacked ...
*(*Former FBI Director James Comey (2014)*)*

The infamous statement from then FBI Director James Comey was limited to companies in the United States being targeted by a specific nation state, but is highly applicable to a wide range of companies today. This is backed by the global information security survey from 2018 by EY [26], which found that only 12 % of the organisations participating in the survey believed they would be able to detect a sophisticated cyber attack. Thus, it is not a question about *if* an organisation is attacked, but rather *when*.

Today's evolving threat environment require organisations to prepare to be compromised. This necessitates dynamic security programs, with continuous improvement of teams, tools and procedures. Organisations need to make sure their IT infrastructure is both *defendable* and *defended*, and have Incident Response Teams (IRTs) in-place to handle intrusions.

1.1 Justification, Motivation and Benefits

One of the most important parts of incident response is also the most often omitted: learning and improving
*(*National Institute of Standards and Technology (NIST) [16, p. 38]*)*

Even though it is noted as one of the most important parts of incident response, organisations are often omitting learning and improving from incidents [16, p. 38]. Interestingly, even though Learning from Incidents (LFI) is emphasised in well-adopted standards and guidelines on incident response [16, 34, 40, 58, 33], several

research papers reveal that organisations lack structured approaches for doing so [56, 35]. This results in organisations missing out on great opportunities to improve their security posture.

Related work suggests that having intelligence capabilities are imperative to ensure efficient LFI [45]. MITRE, which is a none-profit research and development organisation sponsored by the federal government of the United States, argues that it is '*increasingly necessary*' for organisations to have intelligence capabilities [4]. In contrast, the EY survey found that only 47 % of the organisations had a formal intelligence program [26]. Another finding in the EY survey was that only 36 % of the organisations collaborated and shared data with industry peers [26]. This, too, is in collision with the viewpoints of MITRE, which reasons that information sharing with '*partners, peers and others they select to trust*' is paramount for having successful intelligence capabilities [4]. *Those who fail to learn from the mistakes of their predecessors are destined to repeat them*[1]. Thus, learning from your intrusions, and the intrusions of your peers, is vital to improve the security posture of organisations.

This thesis intends to determine how organisations can learn from past intrusions by exploring how industry experts recommends to structure LFI in organisations, and further investigate how this is currently done in two participating organisations.

1.2 Scope and Limitations

This research project explores post-incident activities[2] and how the output of these activities can be used in a holistic manner. Reviewing how the output of post-incident activities could be used in a structured manner is thus within scope, while the monitoring and response part of incident management, including detection, analysis, containment, eradication and the recovery phase, are out of scope. Fig. 1.1 illustrates the relation between post-incident activities and the overall incident response process.

Intrusion analysis itself is on the periphery of the scope, as this thesis mainly focus on what to do with the *outcome* of the analysis, and not the analysis itself. However, it is natural to discuss intrusion analysis and the requirements to the output of such analysis, on a high level.

Additionally, a breakdown of the methodological limitations are given in Sect. 3.7.

[1] Paraphrase of the famous quote by George Santayana (1905): '*Those who cannot remember the past are condemned to repeat it*' [72]

[2] The activities that are conducted *after* an intrusion has been dealt with.

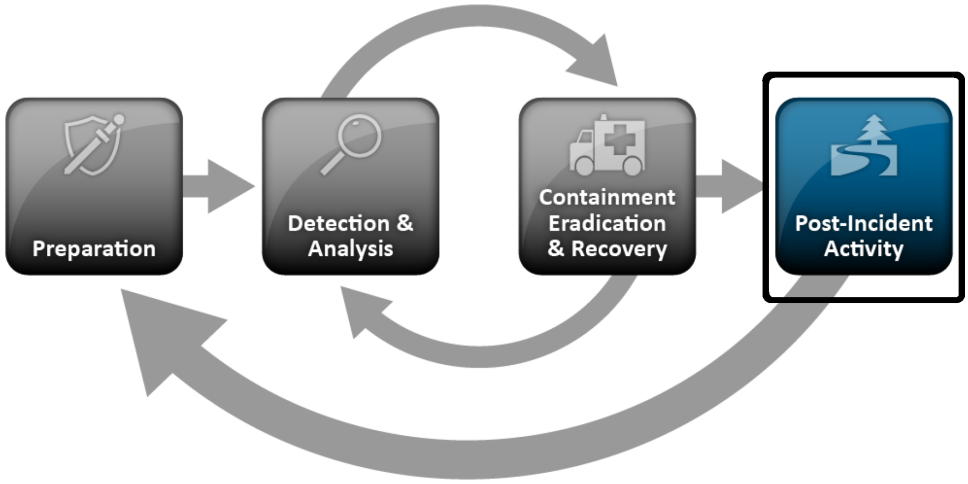

Figure 1.1: Post-Incident Activities in the NIST Incident Response Life Cycle, adapted from Cichonski et al. (2012) [16]. NIST defines four phases of an incident response life cycle; preparation, detection & analysis, containment, eradication & recovery, and post-incident activity [16]. The scope of this thesis are the post-incident activities and how the output from this phase can help improve the security of organisations.

1.3 Contribution

This study offers new insight into how organisations can leverage intrusions to improve their security posture. By analysing information from key experts and describing how organisations operate within the field of incident response and in particular post-incident activities, the thesis has provided new knowledge in a fast-growing field.

Development of an extended incident response model, Intel-Pervaded Incident Response Operations (IPIRO), provides guidance for organisations implementing or structuring learning from intrusions. Combining this with a thorough discussion leveraging intrusions, the thesis can be of significant value for organisations striving to enhance their utilisation of past intrusions. The analysis and model is also of use to researchers studying proactive learning within the information security domain.

1.4 Terminology and Definitions

This thesis will use the following terms as defined in ISO/IEC 27035-1 [34]:

Information security incident: One or multiple related and identified information security events that can harm an organisation's assets or compromise its operations.

Incident Response: Actions taken to mitigate or resolve an information security incident, including those taken to protect and restore the normal operational conditions of an information system and the information stored in it.

Incident Response Team (IRT): Team of appropriately skilled and trusted members of the organisation that handle information security incidents during their lifecycle. **Computer Emergency Response Team (CERT)** and **Computer Security Incident Response Team (CSIRT)** are commonly used terms for **IRT**.

The *Cyber Security Incident Response Guide* by CREST (2013) found that there is no common definition of a *cyber security incident* [18]. As such, we will use the following definitions for a cyber security incident and a cyber security intrusion, simply denoted *incident* and *intrusion* respectively throughout the paper:

Incident[3]**:** An incident caused by a malicious threat actor, that threatens the confidentiality, integrity or availability of information or IT assets

Intrusion: all malicious and nefarious activity targeting computer systems and networks [14]

Incidents are usually given a broader definition, but since this thesis will limit its research to incidents containing an intrusion, it is suitable to use the definition given in the list above.

1.5 Thesis Outline

The following section provides an overview of the structure of the thesis. The structure of the thesis is based on the hourglass approach [12]. The basic idea behind this approach is to

- start with a wide introduction, then
- narrow the scope with a literature review and define the focus of the research based on identified gaps in knowledge, then
- outline a research methodology to answer the research questions, collect data, present results, discuss how the findings relate to the literature and draw conclusions based on this.

[3]Modified based on Frode Hommedal's definition of a cyber security incident

The remaining part of the paper proceeds as follows: Chapter 2 presents a background on intelligence-driven incident response and discusses relevant work. Chapter 3 presents the research method used in this study, discusses the choice of method, and explains how the interviews and a case study were conducted. Chapter 4 presents the empirical data collected using interviews and document studies. Chapter 5 discusses the implications of the research findings, while Chapter 6 presents an extended incident response model to facilitate learning from intrusions. Chapter 7 draws conclusions and provide suggestions for future work.

The method used for selecting bibliography for this section is described in Appendix A, and an overview of standards and guidelines relevant for incident response are presented in Appendix B. Appendix C contains the interview guide used in the interviews.

Chapter 2
Background

This chapter starts with a brief introduction to the *Sliding Scale of Cyber Security* [45] to place *active defence* in context of other security elements. The chapter proceeds to describe the phases of incident response, introduces intelligence, and explains how these two topics are related (Sect. 2.4). This is followed by a presentation and discussion of related research on organisational learning and industrial safety management. The method used for selecting bibliography for this section is described in Apx. A, and an overview of standards and guidelines relevant for incident response are presented in Apx. B.

2.1 Sliding Scale of Cyber Security

The Sliding Scale of Cyber Security is a model of action- and investment categories that organisations can apply to improve their cyber security posture [45]. The motivation behind the model is that organisations would get lower Return on Investment (ROI) on the categories on the right-hand side if issues in the categories on the left are not first addressed. The main objectives of the model are to help individuals and organisations to track and discuss security investments, explain technical details to non-technical individuals and to validate the accuracy of root cause analysis conducted after incidents.

It is important to note that the categories should normally not be equally weighted. Lee (2017) suggests that organisations should structure their resources as a pyramid [49], with 40 % in the architecture category, 30 % in the passive defence category, 20 % in the active defence category, and the last 10 % in the intelligence category. The model, proposed by Lee (2015) in a SANS whitepaper, defines five categories [45]:

1. **Architecture** - '*the planning, establishing, and upkeep of systems with security in mind*'
   Reference architecture models could be, but is not limited to, the National

Institute of Standards and Technology (NIST) 800 Series of special publications, Purdue Enterprise Reference Architecture or the Payment Card Industry Data Security Standard (PCI DSS) [45].

2. **Passive Defence** - '*systems added to the architecture to provide consistent protection against or insight into threats without constant human interaction*' Passive defence models could be, but is not limited to, the NIST 800 Series of special publications, Defence in Depth, and the NIST Cybersecurity Framework [45]. Antivirus (AV) engines, Intrusion Detection System (IDS), Intrusion Prevention System (IPS) and firewalls belong to this category [52, p. 4].

3. **Active Defence** - '*the process of analysts monitoring for, responding to, learning from, and applying their knowledge to threats internal to the network*' The Active Cyber Defense Cycle (ACDC) (Sect. 2.4.5) and Network Security Monitoring are examples of active defence models [45].

4. **Intelligence** - '*the process of collecting data, exploiting it into information, and producing an assessment that satisfies a previously identified knowledge gap*'
   The Kill Chain (Sect. 2.4.1), Diamond Model (Sect. 2.4.1), and Intelligence Life Cycle (Sect. 2.3.3) could be used to structure intelligence processes [45].

5. **Offence** - '*direct action taken against the adversary outside friendly networks ("hack back")*'
   No models for offence are listed since organisations should not be doing offensive operations [45].

The categories are visualised in Fig. 2.1 and illustrates that some actions could fit in the middle of two categories.

2.1.1 Active Defence

When an IT architecture is misconfigured or an adversary is able to circumvent passive defences, active defence plays an indispensable role in detecting and responding to cyber-attacks. Active defence includes incident responders, malware analysts and reverse engineers, threat intelligence operators, members of Security Operations Centers (SOCs), and threat hunters. Active defence should include processes to ensure that the analysts are learning from attackers within their network and that they are equipped to apply the knowledge they gain [45, p. 10].

Industry standards and guidelines on incident response emphasise the importance of having a dedicated team to handle and respond to incidents [16, 34, 40, 58, 33]. Although they use different names for this team, like Incident Response Team (IRT),

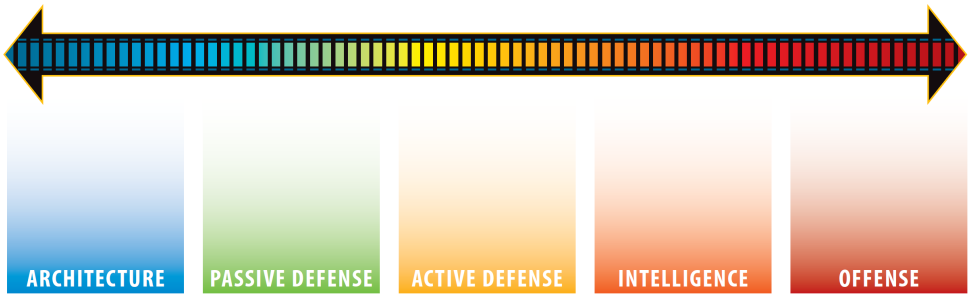

Figure 2.1: The Sliding Scale of Cyber Security, taken from Lee (2015) [45]. Organisations should use the scale to guide their actions and investments, starting on the left-hand side of scale. **Architecture** refers to the design the network and the systems within the network, **passive defence** refers to tools and systems added to the architecture to provide security without depending on human interactions, **active defence** refers to network monitoring, incident response, learning from intrusions and consume intelligence [52, p. 4]. **Intelligence** is both the product and process of producing knowledge from collected data and information to answer specific questions, while **offence** refers to lawful countermeasures that organisations could make for self-defence [52, p. 5].

Computer Emergency Response Team (CERT), and Computer Security Incident Response Team (CSIRT), they all refer to a unit that should be responsible for incident response in an organisation.

Active defence is closely related to intelligence. For instance, the Active Cyber Defense Cycle (ACDC) model has an explicit phase for Cyber Threat Intelligence (CTI) consumption. Analysts responding to incidents would ideally be able to consume intelligence generated by an intelligence unit[45, p. 3], either in-house or outsourced. On the other hand, intelligence is produced by collecting and analysing incident response data which is often generated by active defence. Additionally, CTI is important for the performance of IRTs and can help organisations in prioritising security mechanisms. CTI and intelligence in general, which is further discussed in Sect. 2.3, should enable actions, either on a strategic planning level or as tactical support during an incident response [71, p. 79]. For incident response, CTI is useful for [71, p. 79]:

- Enriching alerts that improve the handling of incidents during the early phases of incidents.
- Enriching and contextualise information found during incident investigations or provided by external parties.

– Informing incident responders and give them a better understanding of the Tactics, Techniques and Procedures (TTPs) associated by the adversaries they are up against.

2.2 Incident Response

Incident response, within the information security domain, relates to preparing for an incident, detecting an incident, analysing and handling the detected incident, recovering from the incident, and finally, using the incident as basis for improving security by learning from it. Although various guidelines and standards describe the overall process of incident response differently, it is common to split the process into a set of phases.

Frequently used industry standards and guidelines on incident response divide incident response into a set of phases [16, 34, 40, 33]. They all use different names and numbers of phases to describe the overall process. However, they all start with an initial preparation phase occurring before incidents take place. Additionally, although *lessons learned* is not always given a separate phase, they all emphasis the need to learn from incidents. ISO/IEC [34] and SANS [40] both have a final phase for *lessons learned*, while NIST [16] and ISF [33] include *lessons learned* as part of their last phase.

The following subsections give an introduction to the life cycle of incident response and the content is, unless specified otherwise, derived from the NIST *Computer Security Incident Handling Guide* [16]. The relationship between the phase are illustrated in Fig. 2.2.

2.2.1 Prepare

This phase involves setting up an IRT, training its team members, deploying systems to help detect and respond to incidents, and identify what is *normal*[1] in an organisation. Implementing mechanisms for risk mitigation is also part of the preparation phase, but the IRT is not necessarily responsible for this part of the preparation.

2.2.2 Detection & Analysis

Organisations should be prepared to detect and handle any incident. Two signals can be used to detect an incident; *precursors*[2] and *indicators*[3]. The presence of a

[1] An analyst or an Intrusion Detection System needs to know what is normal to be able to detect abnormal behaviour in networks and on hosts. This could be baselining network traffic, common application usage or login activities.

[2] *'A precursor is a sign that an incident may occur in the future'* [16]

[3] *'An indicator is a sign that an incident may have occurred or may be occurring now'* [16]

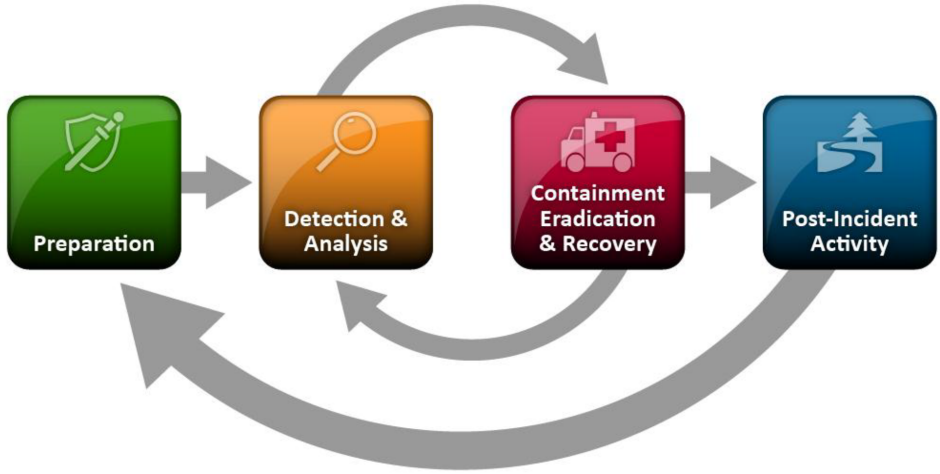

Figure 2.2: Incident Response Life Cycle, taken from Cichonski et al. (2012) [16]

precursor or an indicator will require an analysis to determine if this is in fact an incident, or if it is a false positive. Should it be determined that it is an incident, a policy created in the prepare phase should guide the analyst in giving the incident a suitable classification and criticality level.

2.2.3 Containment, Eradication & Recovery

Predetermined strategies and procedures for containing an incident will ease decision-making related to an incident. A predetermined procedure could be to isolate infected hosts, shut down an infected server and disable compromised user accounts. *Containment* strategies will vary depending on the type of incident, and hence a range of strategies should be predetermined. *Eradication* could be necessary in some incidents, like malware-based incidents or breached user accounts. It is important to identify all affected hosts and user accounts in such incidents and ensure that the eradication procedure is successfully completed before starting the recovery procedure. Both eradication and recovery are system implementation specific, hence few detailed recommendations are included in the reviewed industry standards and guidelines described in Apx. B. *Recovery* consists of getting systems back online, restoring services and enable disabled user accounts.

2.2.4 Post-Incident Activity

After an incident has been dealt with, the organisation should perform some post-incident activities. Creating follow-up reports and conducting *lessons learned* meet-

ings after major incidents are recurring recommendations [16, 34, 40, 58, 33], but a range of other activities are encouraged as well. NIST recommends three activities [16]: *lessons learned*, using collected incident data and, evidence retention. The outcome of these activities should be fed back into the preparation phase to reduce the risk of similar incidents in the future. For instance, if a *lessons learned* meeting reveals that the root cause of the incident was an unpatched server in the Demilitarized Zone (DMZ), then this vulnerability could be mitigated in the preparation phase to avoid it being exploited in the future. *Lessons learned* should determine root causes of an incident, identify missing information during the response which could have improved the handling of the incident, and evaluate if communication and information shared with external parties could have been done better. Collected incident data could be used to aggregate metrics about single incidents. Such metrics could subsequently be used to evaluate IRTs operations, improve risk assessments, and increase security awareness throughout the organisation.

Organisations and sectors should ensure effective coordination and information sharing between, and within, their IRTs and appropriate partners throughout the life cycle of an incident. NIST stresses that it is important to clearly define what type of information should be communicated with partners. Information sharing could be done in an ad hoc manner, where email, instant messaging and phones are used to share information with peers and coordinate strategies for incident response in a cost-effective way. Cross-organisational coordination and information sharing could also be made partially automated by exchanging information in machine readable format. Such an exchange requires that both parties agree upon a common format to structure information.

Root Cause Analysis

Root cause analysis is not a defined process, nor is it a single technique [65]. As Peerally et al. (2017) describes, root cause analysis is '*a range of approaches and tools drawn from fields including human factors and safety science that are used to establish* ***how and why an incident occurred*** *in an attempt to* ***identify how it, and similar problems, might be prevented from happening again***' [65]. For instance, a missing patch on a application server is not the root cause of an intrusion, but rather some patching policy that failed. In this example, the root cause analysis should not be limited in scope to the compromised system only, but investigate why and how a patch could be missing from any system

Peerally et al. (2017) suggests that a problem with root cause analysis is that it is often assumed that the output should be a singular linear cause to the incident. They argue that a root cause analysis should rather identify the factors contributing to the incident taking place, and that there are usually multiple factors contributing

to the incident occurring.

2.3 Intelligence

'Intelligence is what data wants to be when it grows up.'
*(*Bianco (2016) [9]*)*

A common pitfall when talking about intelligence, is to have a conception that data, information and intelligence are different words for the same thing. However, as Fig. 2.3 demonstrates, data, information and intelligence have a hierarchical relationship. To avoid misconceptions about intelligence and what it is, it is imperative to have a common ground of terms and vocabulary. Data, information, knowledge and intelligence are defined below.

Data: '*Data comprise facts, observations, or perceptions (which may or may not be correct). Alone, data represent raw numbers or assertions, and may therefore be devoid of context, meaning, or intent*' [5]

Information: '*Information includes data that possess context, relevance, and purpose. Information typically involves the manipulation or raw data to obtain a more meaningful indication of trends or patterns in the data*' [5]

Knowledge: '*Knowledge in an area is defined as justified beliefs about relationships among concepts relevant to that particular area. Intrinsically different from information*' [5]

In knowledge management the key elements are data, information and knowledge. In this thesis, the terms *data* and *information* will be used further, while *intelligence* will denote the result of the analysis and production of information. In short; intelligence is *'information that has been analysed to answer a specific question' [71], and the associated process of doing so.* Chismon and Ruks (2015) argues that intelligence is actionable information that aids decision makers in making sound decisions [15]. We will use the following definition of intelligence by *U.S. Department of Defense's Joint Publication 2-0: Joint Intelligence* [78] in this thesis:

Intelligence: '*Intelligence is; 1) The product resulting from the collection, processing, integration, evaluation, analysis, and interpretation of available information concerning foreign nations, hostile or potentially hostile forces or elements, or areas of actual or potential operations. 2) The activities that result in the product. 3) The organisations engaged in such activities*' [78]

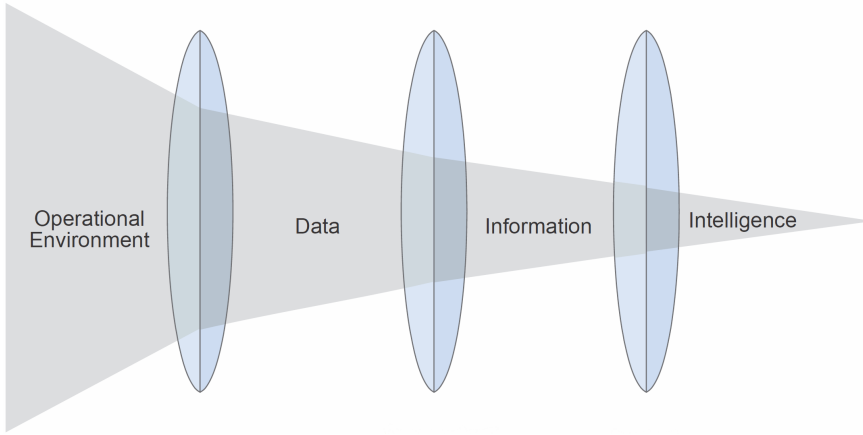

Figure 2.3: The relationship between data, information and intelligence, taken from *the U.S. Department of Defense's Joint Publication 2-0: Joint Intelligence* [78, p. 20]. To have data, you have to collect it. Further, to get information you will need to process and exploit the data. Intelligence is both the process and product of analysing the processed data to answer specific questions.

Another description of intelligence, and its relation to *friendly intelligence* and threat intelligence, is found in a SANS paper about threat hunting [51]:

> '*Intelligence is usable knowledge generated from information. It can be generated about friendly forces (*friendly intelligence*) or about adversaries (*threat intelligence*)*'

Threat intelligence is a subtype of intelligence. This relationship is demonstrated in Fig. 2.4. To be a threat, *'something'* has to have the capability, intent, and opportunity to damage an organisation [47]. Without all those three elements present, *'something'* is not a threat. We will use the following definition of threat intelligence by Gartner [27] in this thesis:

Threat Intelligence: '*evidence-based knowledge – including context, mechanisms, indicators, implications and actionable advice — about an existing or emerging menace or hazard to IT or information assets. It can be used to inform decisions regarding the subject's response to that menace or hazard*' [27]

Cyber Threat Intelligence (CTI) is a further subtype of threat intelligence. There are not any common formal definitions of CTI. In this thesis will we use the following definition by Roberts and Brown (2017) [71]:

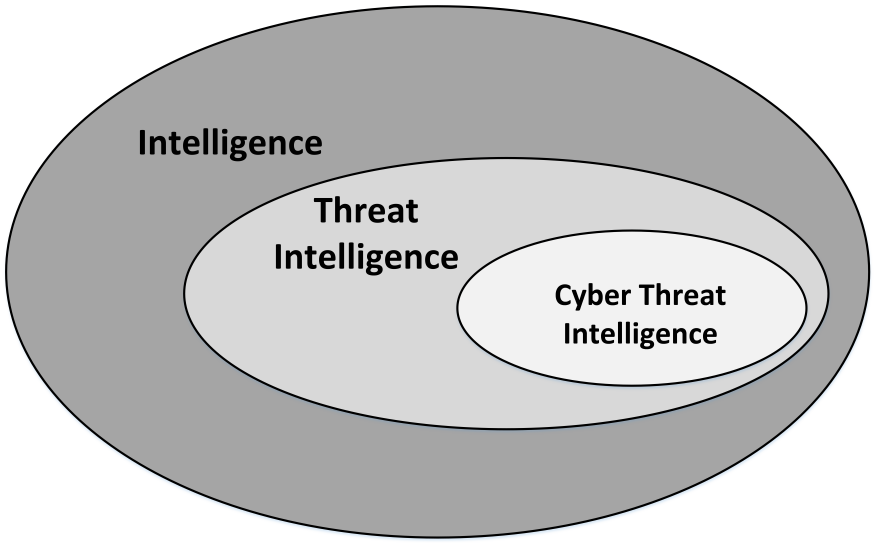

Figure 2.4: Subtypes of Intelligence, adapted from Roberts and Brown (2017) [71, p. 3]. Cyber Threat Intelligence (CTI) is a subtype of Threat Intelligence, which itself is a subtype of Intelligence.

Cyber Threat Intelligence (CTI): '*the analysis of how adversaries use the cyber domain to accomplish their goals.*' [71, p. 3]

Tools cannot generate intelligence [44]. An intelligence analyst will use tools, and the tools will make the analyst more efficient, but it is not possible to take the human element out of the intelligence cycle explained in Sect. 2.3.3.

> *It is important to understand that* ***no tool can produce intelligence. Intelligence is only created by analysts****. The analysis of various sources of information requires [an] understanding [of] the intelligence needs, analysis of competing hypotheses, and subject matter expertise.*
>
> *(*Lee (2015) [44]*)*

2.3.1 Levels of Intelligence

Threat Intelligence and CTI can be found in different levels of abstractions. Chismon and Ruks (2015) defines four levels of Threat Intelligence [15]:

1. **Strategic intelligence:** high-level intelligence consumed by senior decision makers or at board level. Strategic intelligence is naturally found in the form

of prose, which includes written reports, briefings and informal conversations. This type of intelligence does not contain technical details, but could rather focus on the financial impact of cyber activities or attack trends.

2. **Operational:** covers information about specific impending attacks, and is consumed by security managers and incident responders. Chismon and Ruks (2015) notes that this type of intelligence is very rare, and usually only available to government agencies.

3. **Tactical:** contains intelligence on how threat actors behave, what their operations look like and what attack vectors are leveraged. TTPs are often described in intelligence at this level. This level of intelligence is consumed by SOC personnel and incident responders, and is vital to ensure that preventive controls, detection mechanisms and incident response tools and procedures are prepared for the tactics used in their threat environment. Tactical intelligence could be collected by reading industrial white papers, technical blogs, or through informal conversations with peers in the community.

4. **Technical:** low-level data with lower time-to-live and is often exchanged between computers only. Even though Chismon and Ruks (2015) labels this level as '*Technical Threat Intelligence*', it could be argued that this is not threat *intelligence*, but rather threat *data*. Lee (2015) reasons that most threat feeds are not threat intelligence [44]. However, that does not imply that feeds are not useful. Technical threat data would typically include IP addresses, file hashes and domain names. This is often provided without much context other than it is suspected to be related to malicious activity. Missing context to such indicators makes it more challenging for security personnel to act on associated alarms and notifications.

Pyramid of Pain

The *Pyramid of Pain* depicts the relationship between types of indicators and the associated 'pain' for adversaries to change them [7]. The model is illustrated in Fig. 2.5. The higher levels in the model are more central to an actor's tool chain and objectives [71], and are thus harder to change. The lower levels in the model, on the other hand, are much easier for an actor to change. An implication of this is that the life-span of an indicator will depend on where in the pyramid it is located. The higher up; the longer expected life-span. Indicators from the higher levels are often found in tactical threat intelligence, while the lower levels are found in technical threat feeds.

- **Hash Values**
  Hash values are cryptographic hashes of specific suspected or verified malicious

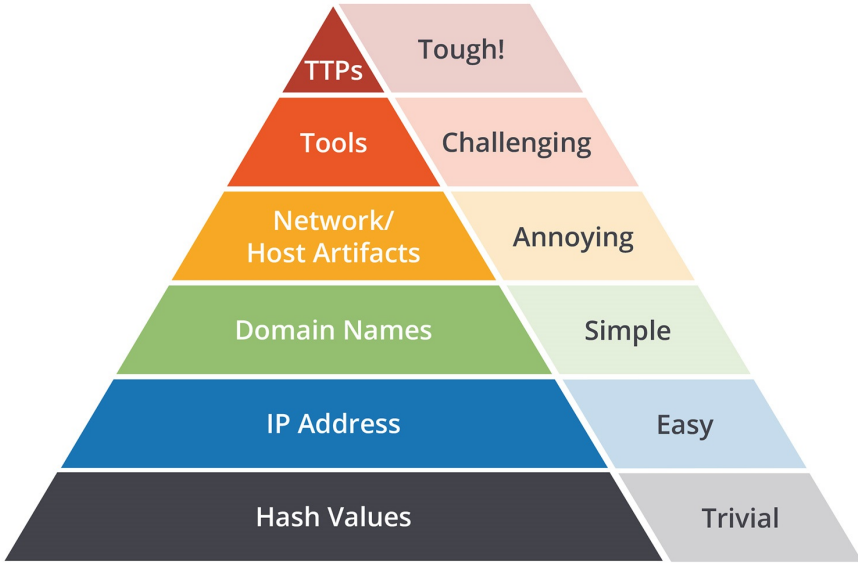

Figure 2.5: The Pyramid of Pain [7], taken from Pace (2016) [62]. The level of difficulty for adversaries to change their indicators increases as you go from low to high. Changing the hash of a file is trivial, while substituting a tool or rebuilding an operation infrastructure takes more resources.

files. The cryptographic hashes could be produced with functions such as SHA1, MD5 or similar cryptographic hash functions. Changing one bit in a file would completely change the corresponding hash value of the new file.

- **IP Addresses**
  IP addresses or IP address ranges that are associated with malicious activity. An IP address could for instance be associated with a Command and Control (C2) server used by a threat actor.

- **Domain Names**
  Domain names, either the domain itself or a sub-domain, which are associated with malicious activity. As with IP Addresses, domain names could be associated with a C2 server. They are slightly harder for adversaries to change, since they need to be registered and propagated throughout the internet.

- **Network and Host Artefacts**
  Network artefacts are observables caused by adversary network activity. This could for instance be URI patterns, HTTP User-Agents, mail headers or C2-traffic patterns. Host artefacts are observables caused by adversary host activity. This could for instance be modified registry keys or dropped file names.

- **Tools**
  Tools are the software used by adversaries to achieve objectives and complete missions. The tools could for instance be used to create malicious documents, dump credentials on a host, establish C2-communication or crack passwords within the victim infrastructure. Most of the tools will be software that are loaded on the victim system post incident, but could be software running remotely as well. Learning how to detect a tool used by a threat actor could force the actor to either build a new tool or learn how to achieve the same goals using a different tool.

- **Tactics, Techniques and Procedures (TTPs)**
  TTPs covers all actions taken by an adversary to achieve some objectives, also known as behavioural indicators [71, p. 79]. It covers the entire kill chain, from reconnaissance and delivery to C2 and actions on objectives. TTPs describe the behaviour of adversaries, and are very challenging to change. It could range from crafting spear-phishing emails with malicious macros inside word-documents to achieve a foothold in a victim's infrastructure. Other TTPs could be to compromise legitimate websites, collect NTLM-hashes[4] from the visiting users, crack the hashes and use credentials to access victim networks via a VPN connection.

2.3.2 Sharing Intelligence

As will be further explained in Sect. 2.3.3, generated intelligence could be shared with stakeholders as part of a dissemination process. Within intelligence, you have to walk a fine line between the 'need-to-know'-principle and the 'need-to-share'-principle [15]. Chismon and Ruks (2015) argues that all levels of threat intelligence could help organisations defend against attacks [15]. Since many attacks are part of an attack campaign, receiving CTI about ongoing attacks from peer organisations could enable an organisation to initiate proactive measures to defend against the attack campaign before they are hit themselves.

Sharing intelligence requires *trust* between the sharing parties [15]. The sharing organisation needs to be assured that the receiving organisations will treat the intelligence with care to keep it from both adversaries and the general public. Therefore, closed and trusted sharing groups are established to facilitate sharing with trusted members [15]. Chismon and Ruks (2015) points out that sharing within trusted personal relationships with peers in similar roles in other organisations often is among the most useful ways of sharing intelligence.

[4] A *NT LAN Manager* (NTLM) hash is used in to authenticate users in windows environments

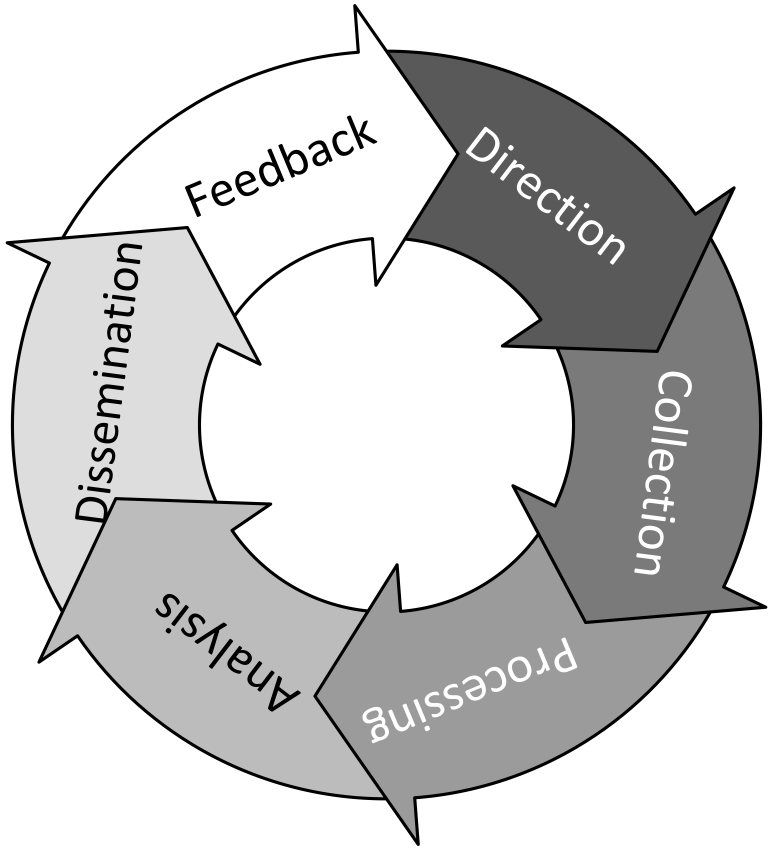

Figure 2.6: The intelligence cycle, adapted from Roberts (2015) [69]

2.3.3 Intelligence Cycle

The intelligence cycle can be divided into a different number of phases. Fig. 2.6 demonstrates the cyclic relationships between the phases. In this section, we will describe an intelligence cycle with six phases, as suggested by Roberts (2015) [69]:

1) Direction

This phase sets the parameters of which questions should be answered. This could for instance be questions related to an Advanced Persistent Threat (APT)[5] group's goals, or what Indicators of Compromise (IOCs) that could

[5]Axiom 7 from the Diamond Model of Intrusion Analysis is borrowed to define an Advanced Persistent Threat (APT) group: '*There exists a sub-set of the set of adversaries which have the motivation, resources, and capabilities to sustain malicious effects for a significant length of time against one or more victims while resisting mitigation efforts. Adversary-Victim relationships in this sub-set are called* ***persistent adversary relationships***' [14, p. 21].

identify an APT group.

2) Collection:
The collection phase involves gathering information required to answer the questions from the first phase. The collection is typically iterative; each new piece of information you collect could trigger collection using new search terms or sources. The collection phase ends when you are out of time, or you cannot find any new relevant data.

3) Processing:
Before the data can be analysed it has to be transformed to formats suitable for analysis. This is done during the processing phase. This processing should ensure consistent analysis and makes it easier to work on the data in the next phase. The data could for instance be formatted as JSON or XML, and ingested into a threat intelligence platform.

4) Analysis:
The analysis phase should focus on answering the questions from the direction phase using the collected data. Written 'long form' reports could be suitable to answer questions about the goals of APT groups, while indicators of their activities might be found by reversing a malware sample. Details of the analysis process going from data and information to intelligence is out of scope for this thesis.

5) Dissemination:
The analysis should be shared in a format suitable for the receiver. There might not be a single format that fits the requirements for all stakeholders. A security engineer could prefer machine-readable formats like JSON and XML, while a summary PDF report might be more suitable for managers and senior management. In addition to finding the right format to share the information, the right level of information in terms of operational security should be considered during this phase. This includes labelling the shared intelligence with the correct sensitivity label[6], and ensure that the receiving parties understand and conform to the inherent label.

6) Feedback:
When the generated intelligence has been shared with stakeholders, an evaluation of the process should be conducted. Were the questions answered? Were the findings presented in a way that the stakeholders could perceive and understand them? The feedback phase is an explicit step to learn from the cycle in order to improve future iterations.

[6]Could be a Traffic Light Protocol (TLP) colour

The quality of generated intelligence is mainly determined by two factors: the collection sources and how the information has been analysed [71, p. 23].

2.4 Intelligence-Driven Incident Response

2.4.1 Intrusion Analysis

Analysing an intrusion, when applicable, is an important part of the post-incident activities of an incident. Two intrusion analysis models that have caught much attention the last decade are the *Kill Chain* and the *Diamond Model of Intrusion Analysis*, which are both described below. Roberts and Brown (2017) point out that these two models can complement each other to give more depth when analysing intrusions [71].

Kill Chain

Kill chains have been used for decades in military threat intelligence and were made mainstream in information security by a white paper labelled '*Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains*' by Hutchins et al. (2011) [32]. The kill chain[7] described in that paper is shown in Fig. 2.7. The purpose of the kill chain is to **describe the sequence of steps an attacker must do to achieve an objective**. It could be regarded as the life cycle of an intrusion from the attacker's point of view. The kill chain model could be applied during the post-incident activities to describe the incident by abstracting the attacker's TTPs. Indicators and information found during intrusion analysis could be structured according to this model. This information could then be used to group similar incidents based on similarities between the associated kill chains, improve defences by intercepting future attacks in the early stages, and aid in visualising incident information to improve security awareness in organisations.

We will use the term *backwards analysis*[8] for analysis focusing on the steps in the kill chain occurring *preceding* the disruption of the intrusion, and *forwards analysis*[9] for the deductive reasoning of what could have happened in the *following* stages. The latter type of analysis is especially applicable in unsuccessful intrusions. The two types of analysis are illustrated in Fig. 2.8

[7] Also called the cyber kill chain when used in an information security setting, but for simplicity we will simply use the term kill chain in this report

[8] Termed *analysis* in the Kill Chain proposed by Hutchins et al. (2011) [32]

[9] Termed *synthesis* in the Kill Chain proposed by Hutchins et al. (2011) [32]

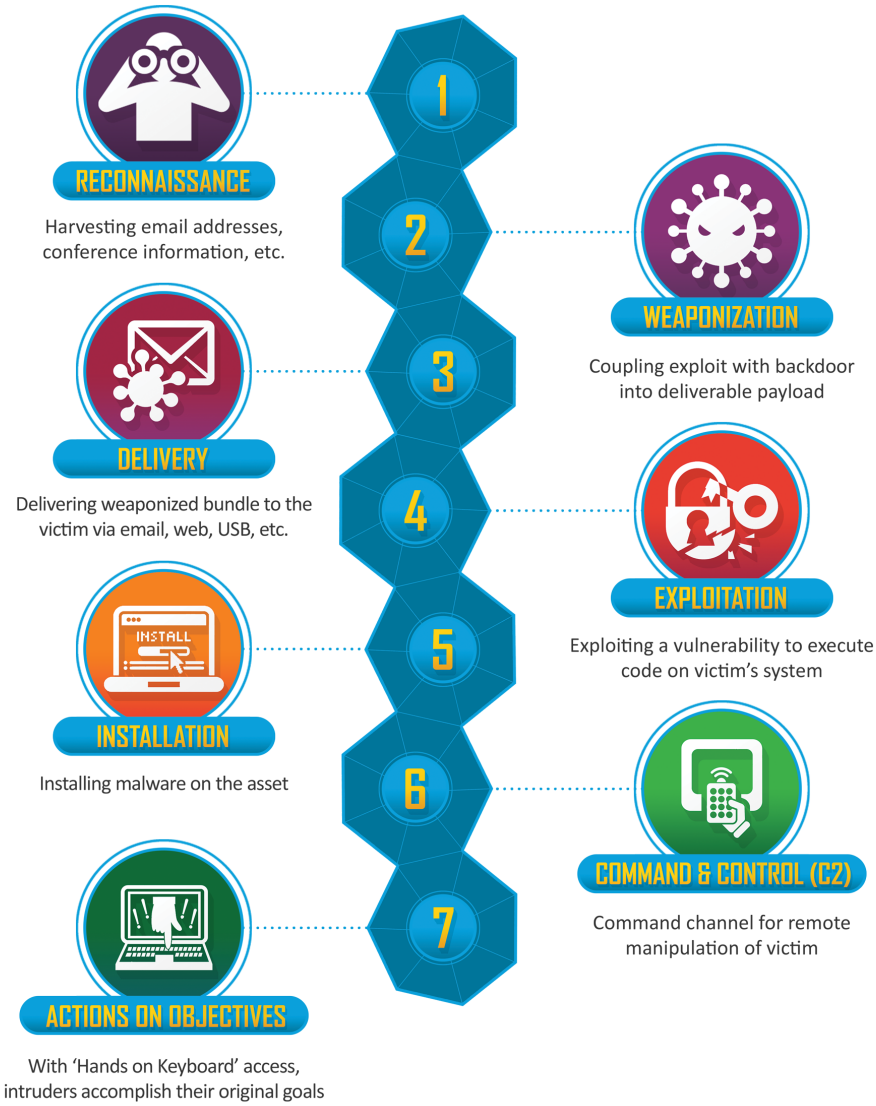

Figure 2.7: The kill chain [32], taken from [57]. The *delivery* stage is the first stage where the attacker must be active and interact with the victim [71]. As with the incident response life cycle defined by NIST, described in Sect. B.1, the kill chain could be looped such that the objective of the first kill chain is to do reconnaissance for the second kill chain. The second kill chain could then have the final objective, or it could be another step in an attack.

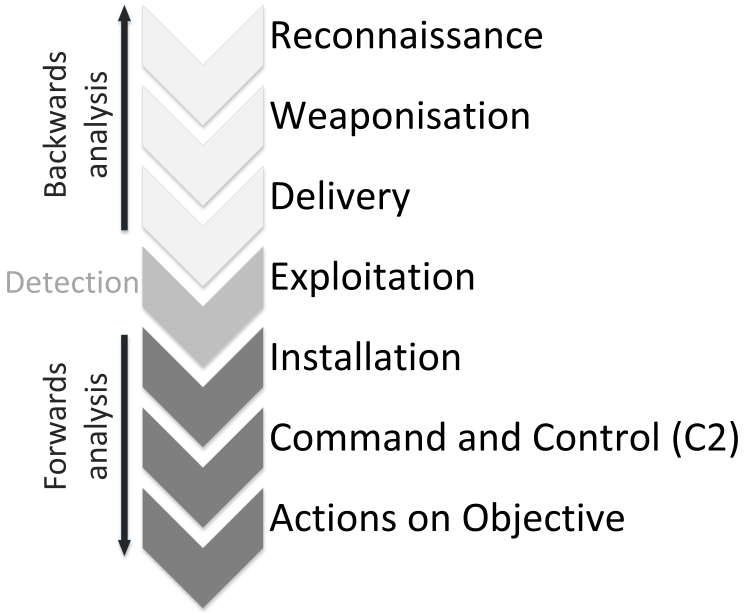

Figure 2.8: Forward and backward analysis with the kill chain, adapted from Hutchins et al. (2011) [32]

Diamond Model of Intrusion Analysis

The Diamond Model of Intrusion Analysis describes malicious events in an incident with four connected features; adversary, infrastructure, capability, and victim. These core features are connected by '*an* ***adversary*** *deploying a* ***capability*** *over some* ***infrastructure*** *against a* ***victim***' [14, p. 7]. The model, illustrated in Fig. 2.9, may seem trivial at first sight, but has proved to be powerful and applicable in a range of incident investigations [48]. The diamond model and the kill chain can complement each other by categorising each event in the diamond model according to its stage in the kill chain. This will enrich the information in the kill chain and give extended context to events in the diamond model, and could be used to infer if an incident is part of a larger attack campaign. The diamond model could also be used in combination with the kill chain to help produce threat intelligence by extracting indicators and information from interactions with an adversary [45, p. 15].

2.4.2 Storing and Sharing Information

Incident response actions should be analysed throughout the life cycle of incidents, from the detection phase to the post-incident activities, and could be stored on a format which makes it machine readable, simple for a human analyst to use, and

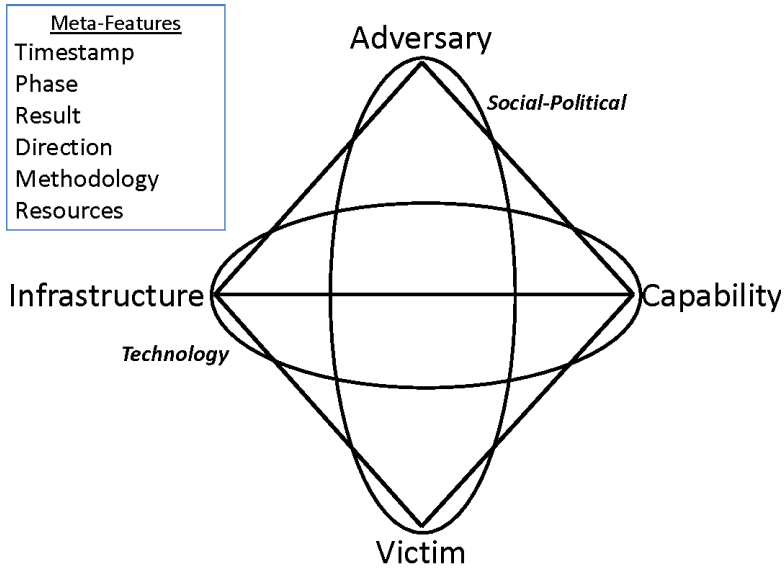

Figure 2.9: The Diamond Model of Intrusion Analysis, taken from Caltagirone (2013) [14]. Events are represented by four core features, namely *adversary, victim, capability* and *infrastructure*. These four features are, according to Caltagirone et al. (2013), present in all malicious activities [14, 13]. The meta-features listed to the left of the diamond represents important, but non-critical, features of events [14, p. 15]. A social-political relation between the adversary and the victim(s) are always present [14, p. 20], in addition to a technological relation between some capabilities deployed over an infrastructure [14, p. 24]

be adoptable for a wide range of incident types. *Structured Threat Information eXpression (STIX)* [4] is an open source project supported by Oasis to store technical and tactical information. It is a language and serialisation format that can be used to exchange CTI. It consists of twelve different object types and two relationship types to connect the objects. STIX allows for contextual details in objects and relationships. *Trusted Automated eXchange of Indicator Information (TAXII)* [39] is a transportation and sharing framework, and is often paired with STIX. It defines methods for sharing CTI between entities and organisations. *Vocabulary for Event Recording and Incident Sharing (VERIS)* facilitates storage of strategic information that organisations can use to understand risk environment they operate in [71, p. 131]. VERIS is a good fit to simplify and share complex incident information. It includes metrics about an incident, like response time, containment time and severity of the impact. [71, p. 131]. This makes VERIS a good fit for evaluating IRTs and to provide strategic information to security management. STIX, on the other hand, is more suitable for generating rules and alerts in an IDS/IPS, and its ability to store TTPs information could be exploited in threat hunting.

2.4.3 Threat Hunting

Threat hunting is a proactive approach to detect threats, and falls in the *Active Defence* category on the Sliding Scale of Cyber Security [51]. SANS defines *threat hunting* as a *«proactive and iterative approach to detect threats»* [51]. A hunt starts with the hunter forming a testable idea of what threats might be in her environment and how such threats could be found [51]. Such testable ideas are referred to as threat hunting hypotheses, and they are key to the success of any hunts. It is important that a hypothesis is formulated such that it is possible to verify it with the data available to the hunter. A hunter's knowledge about the data sources available, and the tools and technologies required to explore the data, is vital for the success of a hunt [51]. If the activity that the hunter is looking for simply is not present in the network, there is nothing for her to find. Yet the hunter could still identify detection or logging gaps in the network that should be fixed. As such, a hunt could be successful even if it did not detect any malicious activity. One of the main goals of threat hunting should be to improve automated detection capabilities by turning a hunt into codified detection logic.

Threat Hunting Hypothesis

A threat hunting hypothesis should answer four questions [13]: 1) what should be hunted?, 2) where could it be found?, 3) how could it be found?, and 4) how long should the hunt last? Answers to these questions ensure that the hunter has scoped the hunt in terms of objective and data sources, identified tools to assist her, and available time. Without the last one, a threat hunter could end up in a never-ending chase for a threat that simply is not present in the environment.

There are mainly three sources that hypotheses can be rooted from, and a hypothesis can be derived as a combination of multiple of these sources [51]:

- Friendly or threat intelligence
- Situational awareness
- Domain expertise

Three types of threat hunting hypotheses are explored below, each type stemming from one of the sources listed above. The content about the different types of threat hunting hypotheses is, unless specified otherwise, derived from [51].

Intelligence-Driven Hypotheses

Refined and contextualised threat intelligence could be used to generate threat hunting hypotheses. This could provide quick discoveries of threats in an environment if

the hunter has access to relevant intelligence. Knowledge about adversaries' IOCs and TTPs could be used as a starting point when generating hypotheses, but it is important to have context around IOCs for them to provide valuable input. Bad IOCs will give high amounts of false-positives, and without the context of what phase of an attack an IOC is related to, it is difficult for a hunter to collect the right data [51]. A hunter should strive to climb the Pyramid of Pain [7] by hunting for TTPs rather than atomic IOCs. Regardless, IOCs could give results that could guide a hunter in prioritising which alarms and notifications that should be investigated. Further, how adversaries *use and leave* IOCs in the environment could itself be hunted for. This could for instance be how C2 traffic is obfuscated, or how IOCs are overlapping between multiple attack campaigns.

Situational awareness
Knowing what elements are on an organisation's network, and which elements that are *not*, is vital to be able to detect changes in the network. Not knowing what the network topology normally looks like makes it difficult to detect changes[10]. A hunter could gain situational awareness from friendly intelligence which should provide an understanding of the organisation's network and business environment. Threat hunters having situational awareness could generate hypotheses regarding adversary activity that could happen within their network and avoid spending time and resources on hypotheses that focus on data or technologies that are not present in their network.

Domain expertise
The experience, background and skills of a hunter all influence the hunt and the generated hypothesis for the hunt [51]. The authors of '*Generating Hypotheses for Successful Threat Hunting*' argue that [51]:

> '*[...] a hunter's previous hunts and engagements with adversaries influence later hypotheses, even for unrelated threats in new environments. [...] domain expertise is the combination of situational awareness and intelligence-driven understanding in a historical context*'

Thus, having good domain expertise implies having the prerequisite knowledge about both the threat landscape and the environment that the hunter's organisation is operating within. Such expertise is useful when generating hypotheses and finding the right data sources to answer the questions raised in the hypotheses. It should be noted that bias is often an unwanted side effect of experience, and it is, as such, imperative that the hunter is aware of this in order to defeat her cognitive biases [51].

[10]Changes in topology in traditional IT environments can be difficult to use in hypotheses due to the dynamic nature of such environments. However, Industrial Control Systems (ICS) networks are naturally more static, and thus changes to the network topology is rare [43].

The Hunting Maturity Model

Although it is common that threat hunting occurs ad hoc in organisations, it is not straight forward to integrate threat hunting into a holistic security program [52, p. 3]. The level at which an organisation is able to integrate threat hunting into their security work-flows depends on their security maturity as an organisation.

> *'[. . .] threat hunting is accessible to all, but an organisation must be mature enough to get a proper return on investment from it and make it a repeatable and consistent process'* (Lee and Lee (2016) [52])

The Hunting Maturity Model, a model developed by Bianco (2015) to measure an organisation's ability to proactively discover intrusions, lists three factors defining an organisation's ability to hunt [8]:

- Quality of the data available to the hunter
- Tools available to collect and analyse the data
- Skills of the threat hunter

Threat hunting is part of the *active defence* category in terms of the *Sliding Scale of Cybersecurity* [45] and an organisation's ability to hunt is closely related to its *architecture* and *passive defence* maturity [52, p. 5]. For instance, it is difficult to hunt threat actors if the architecture does not support wide ranging telemetry [67]. Similarly, an architecture full of vulnerabilities and passive defences not tuned to the organisation could make hunting challenging, as one could expect the network inhibit noisy commodity malware [52, p. 5]. Thus, for an organisation to move up the Hunting Maturity Model, it should start with the architecture and passive defences to ensure threat hunters are able to conduct effective and efficient hunts.

The Hunting Maturity Model uses five levels to rate organisations' ability to hunt, with the levels increasing in sophistication [8]:

1. **Initial** – automated tools, like IDS and AV engines, are used to detect intrusions. Threat data feeds of technical indicators or signatures may be consumed by these tools. The activities described at this level are not regarded as threat hunting.
2. **Minimal** – automated tools, like the previous level, are used to detect intrusions. Detection is based on threat intelligence consumption from both open and closed sources. At this level, organisations are able to extract technical IOCs

from threat intelligence reports to search historical data. Searching historical data with technical IOCs is a very basic form of threat hunting.

3. **Procedural** – threat hunting is part of the security program at organisations on this level, with hunters being able to learn and implement procedures developed by others. This is arguably the most common maturity level among organisations that have threat hunting programs.

4. **Innovative** – at this level, instead of relying on hypotheses and procedures developed by others, organisations are developing and publishing their own procedures. Additionally, hunts are documented and repeated frequently.

5. **Leading** – at this level, all hunts end up as automated detection logic. Rather than repeating a hunt multiple times, the hunt is operationalised and automated to allow the hunter to move on to the next hypothesis.

2.4.4 F3EAD

Find, Fix, Finish, Exploit, Analyse and Disseminate (F3EAD) is originally a methodology for combining kinetic operations with the intelligence cycle [68], and was designed and adapted for the U.S Foreign Internal Defence missions in Latin America in the 1980s [28]. Find, Fix, Finish, Exploit, Analyse and Disseminate (F3EAD) is applicable to cyber security as well, as it combines security operations with the intelligence cycle. The first three phases (Find, Fix, Finish) are the security operation phases, while the latter three phases (Exploit, Analyse, Disseminate) are the intelligence phases. Fig. 2.10 demonstrates the cyclic nature of the F3EAD model.

Find

The first phase of the F3EAD process is *Find*. In this phase it is decided *what the focus of the operation should be* [68]. In terms of the incident response life cycle, it is associated with the prepare phase [71, p. 54]. It could for instance be that a threat intelligence vendor releases a report about a new APT group, or some new TTPs they have associated with an APT group. The focus of the operation is thus the APT group mentioned in the report. Another common targeting methodology in immature organisations is what Roberts and Brown (2017) call *News-Centric Targeting* [71]. This occurs when an executive sees or hears something on public news, and asks the threat intelligence team to analyse the implications of this threat. In such cases, if a news article describes an active APT group, this APT group would be the focus of the operation.

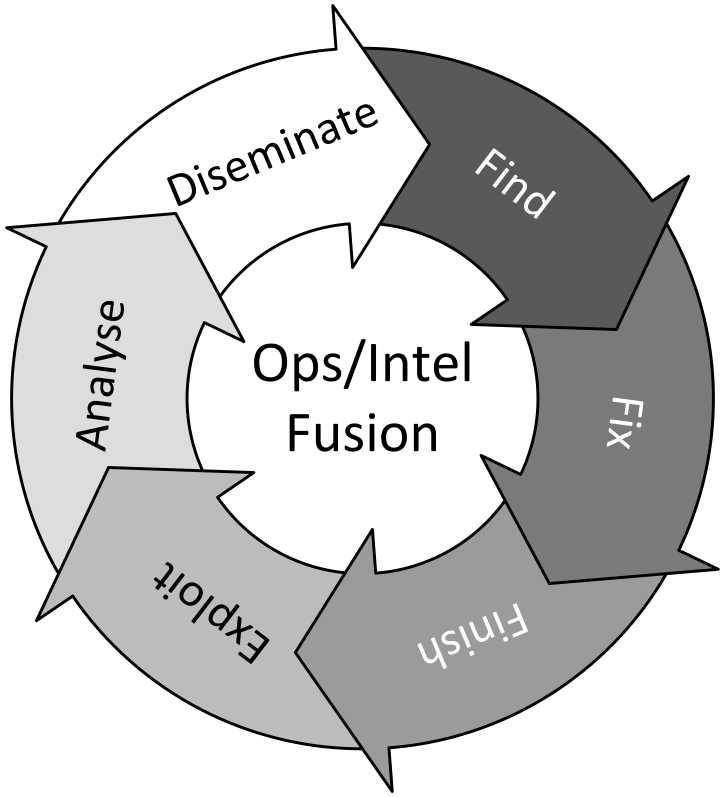

Figure 2.10: Find, Fix, Finish, Exploit, Analyse and Disseminate (F3EAD), adopted from Roberts (2015) [68]. The security operations feed into intelligence from Finish to Exploit, while intelligence informs security operations from Disseminate to Find.

Fix

The name *fix* is somewhat misleading, as noted by Roberts and Brown (2017), because the phase does not involve repairing anything [71, p. 54]. Rather, the fix phase identifies operational presence of the adversary on the victim's network. In terms of the incident response life cycle, it is associated with the identification phase [71, p. 54]. There are many ways to identify the location of adversaries, including their TTPs, mission goals, and using IOCs [71, p. 79]. Malware analysis, disk analysis, network analysis, intrusion investigation and live response are common reactive incident response activities during the fix phase. Proactive activities, such as threat hunting, are applicable during this phase as well.

Finish

The finish phase completes the security operation. An operation is finished when the predetermined objectives from the *find* phase have been met [68]. This includes doing eradication, containment and recovery. Therefore, this phase maps to the third phase in NIST standard on incident response [16]. The actions taken to remove an adversary from the network could be to *detect, deny, disrupt, degrade, deceive or destroy* the actions of the adversary. These actions, known as the *Courses of Action* [38, p. 9], should be taken against an adversary *within* its own network. The *finish* phase nor the Courses of Action imply 'hack-back' activities.

After the *Finish* phase, a transition into the intelligence phases of F3EAD happens, where information collected throughout the first three phases are passed on to the following intelligence phases.

Exploit

During the exploit phase, the information passed on from the previous phases are collected and enriched. This involves storing the information in a format where it can be analysed and used in subsequent iterations of the F3EAD-cycle. The formats discussed in Sect. 2.4.2, such as VERIS, STIX and TAXII, are relevant formats to store and share threat information and incident response artefacts. A threat intelligence platform, which is simply a database and user interface designed to store and handle threat information, could be used to simplify these tasks. The *exploit* phase maps to the *collection* phase in the intelligence cycle.

Analyse

The analyse phase in the F3EAD model is the same as the *analyse* phase in the intelligence cycle described in Sect. 2.3.3.

Disseminate

The analyse phase in the F3EAD model is also very similar to the *disseminate* phase in the intelligence cycle described in Sect. 2.3.3. The difference between the two phases is that disseminate in the F3EAD model informs incident response explicitly via the *find* phase, while the disseminate phase in the intelligence cycle has an explicit feedback phase first. However, IRTs are common receivers of intelligence in the intelligence cycle, and teams other than IRTs are common stakeholders in the F3EAD model.

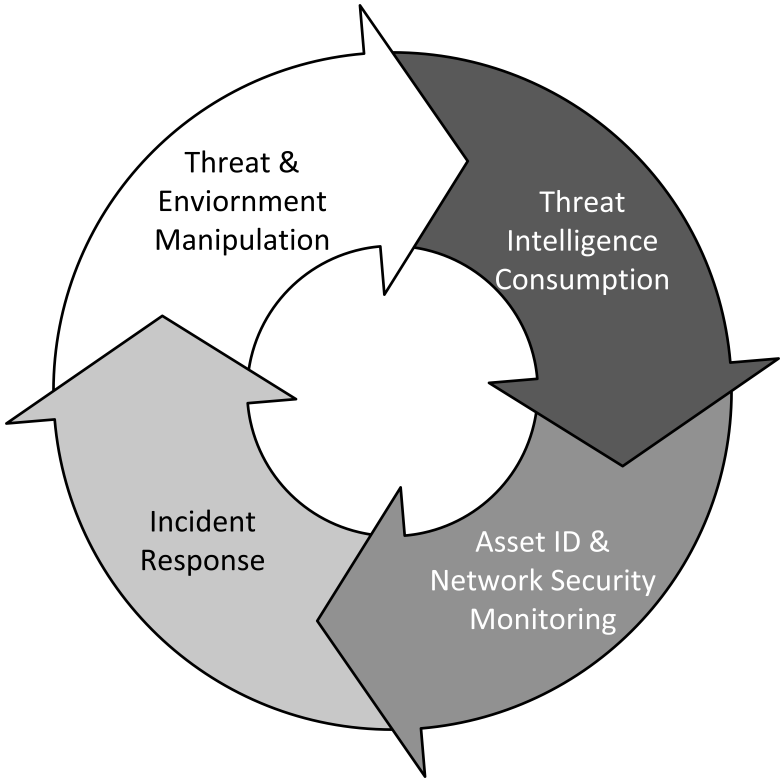

Figure 2.11: Active Cyber Defence Cycle (ACDC), adapted from Lee (2015) [42]

2.4.5 ACDC

The Active Cyber Defense Cycle (ACDC) is a model incorporating intelligence consumption into security operations [46]. The cycle consists of four phases. The relationship between the phases is demonstrated in Fig. 2.11. Before implementing the Active Cyber Defense Cycle (ACDC) model, an organisation should be on top of their architecture and passive defences, as discussed in Sect. 2.1. Doing active defence implies that security personnel is actively handling and responding to threats. For them to be able do their job effectively and efficient, the organisation should have an architecture that is *defendable* [45]. Passive defences are required to increase security in the organisation without requiring active involvement from analysts [46].

Threat Intelligence Consumption

During the *Threat Intelligence Consumption* phase, the analysts should identify intelligence sources and data feeds, both internally and externally, that are related to the *threat landscape* that their organisation operates within [47]. The *threat landscape*

is based on an understanding of the organisation's environment and business mission. The intelligence consumed and filtered by the threat intelligence team should be shared with other members of the security teams. With ACDC, the threat intelligence team works together with other security teams to identify and select the right data and intelligence sources ahead of time, and to ensure that the consumed intelligence is relevant for the organisation's business [42]

As Fig 2.11 demonstrates, the team responsible for network security monitoring would be the first to receive this intelligence. A full iteration of the intelligence cycle is conducted during this phase. The direction phase should end with the analysts having a thorough understanding of the organisation so that they can ask the right questions. As described above, intelligence and data are collected and filtered[11] before being disseminated to the stakeholders which are mainly other members of the active defence team.

Asset Identification and Network Security Monitoring

In the *Asset Identification and Network Security Monitoring* phase, the analysts should ensure that they know what assets need to be secured and what normal operations look like [43]. Identifying assets in ICS networks are often easier due to the static nature of such operations, making asset identification easier [42]. This can be more challenging in traditional IT networks, hence the requirement for proper architecture and passive defences to ease active defence operations [45]. Identifying assets is commonly done with physical inspection, configuration file analysis, and passive and active scanning [43]. Some of these approached could be unsuitable in some environments. For instance, active scanning should not be conducted in Supervisory Control And Data Acquisition (SCADA) networks with legacy equipment, as it could result in unscheduled operation downtime. Passive scanning often has a good ROI with low risk, as it can be used to detect communicating devices at central networking nodes without direct interactions with the communicating devices [43].

Networks should be monitored to detect malicious activity. Network monitoring analysts use IOCs and TTPs provided by the threat intelligence team to identify threats in their network. Taking a proactive role in network security monitoring involves three steps: 1) Collection, 2) Detection, and 3) Analysis. A thorough understanding of the network is required to know which data to collect. Once data is collected, it is analysed to detect malicious activity. Once suspected malicious activity is detected, it has to be analysed in order to verify the threat. Activities that look malicious could be benign, or could look malicious due to a gap in the analyst's knowledge about normal network flows and topologies [43]. If the activity

[11] In the ACDC model, intelligence is consumed rather than generated. Thus, the analyse phase of the intelligence cycle is reduced to filtering out which intelligence is relevant rather than doing a full-scale analysis as described in Sect. 2.3.3.

is verified to be malicious, or strongly believed to be so, an incident response process is initiated.

Incident Response

The *Incident Response* phase of the ACDC model covers the traditional life cycle of incident response, as described in Sect. 2.2. It should be pointed out that even though the *Threat Intelligence Consumption* phase in Fig. 2.11 only seems to feed into the *Asset Identification and Network Security Monitoring* phase, the intelligence is further passed through to the *Incident Response* phase. IRTs are responsible of collecting artefacts and information about threat actors. This is then fed into the next phase; *Threat and Environment Manipulation.*

Threat and Environment Manipulation

During the *Threat and Environment Manipulation* phase, analysts exploit a thorough understanding of the threat to work with teams responsible for the architecture and passive defences in the organisation to reduce the effectiveness of the threat [47]. Having a thorough understanding of the threat requires quality data handed over from the IRT, and tools and resources to enrich that data. Associated malware should be reverse engineered to fully understand the capabilities of the malware [47]. Reducing the effectiveness of a threat varies depending on its capabilities, but could involve writing YARA-rules that block certain traffic or routing traffic to specific IP addresses and domains to a sinkhole server.

The *Threat and Environment Manipulation* phase should inform the *Threat Intelligence Consumption* phase about any IOCs or threat data that they have gained by learning from the threat [42]. The threat intelligence consumption team could then combine this information with external intelligence, which would give the Network Security Monitoring team a better starting point for detecting new malicious activity.

2.5 Related Research

So far this chapter has focused on incident response and intelligence. This section provides a brief discussion on a sample of available academic literature addressing incident learning systems. The literature below has been found and selected based on the method described in Apx. A.

To date, several studies have investigated how organisations can learn from safety incidents and how organisations should structure such learning. We argue that findings and results from studies focusing on *safety* incidents in relation to Learning from Incidents (LFI) are applicable to *security* incidents as well. Relevant literature

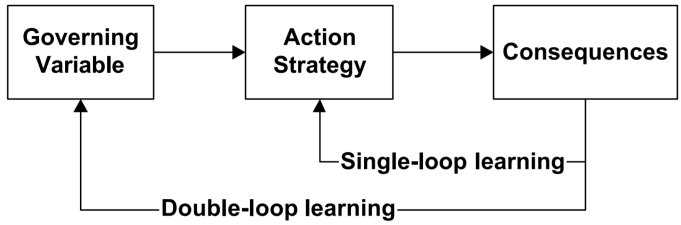

Figure 2.12: Single and double-loop learning for incident response, taken from van Niekerk and von Solms (2004) [79]. Double-loop learning acts on the governing variables in an organisation, while single-loop learning acts on the action strategies.

from other fields are research within the area of organisational learning and industrial safety management.

Duncan and Weiss (1979) defines organisational learning to be '*the process within the organisation by which knowledge about action-outcome relationships and the effect of the environment on these relationships is developed*' [25]. Argyris and Schön (1974) defined the three elements of a learning process to be [3, 79]:

1. **Governing variables:** '*Those dimensions that people are trying to keep within acceptable limits*'
2. **Action strategies:** '*The moves and plans people use to keep the governing variables within the acceptable range*'
3. **Consequences:** '*What happens as the result of an action*'

Van Niekerk and von Solms (2004) argue that the governing variables could refer to acceptable levels of risk, an action strategy could be the procedures outlining accepted employee behaviour in specific scenarios, while consequences would be both intended and unintended results [79]. If the underlying governing variables for a system are taken for granted, *single-loop learning* is often present in the system. Single-loop learning would only impose specific actions in response to identified issues [73]. *Double-loop learning*, on the other hand, could impose changes to the governing variables setting the directions for the action strategies [79]. Malhotra (2006) reasoned that double-loop learning is needed to achieve efficient organisational learning [59]. Fig. 2.12 illustrates the relationships between the three elements and the two types of learning. Drupsteen and Guldenmund (2014) found that difficulties in identifying the organisational factors and managerial weakness that enabled incidents to happen often caused learning after incidents to be limited to single-loop learning, where only the direct causes are addressed [24].

Shedden et al. (2010) found that effective incident learning could increase an organisation's ability to manage their incident response capability, make improvements, and communicate learning notes to stakeholders [73]. To facilitate such organisational learning, they suggested that organisations should incorporate double-loop learning into their incident learning activities, agile incident learning feedback to avoid knowledge erosion, and a holistic dissemination process ensuring that knowledge is transferred to all relevant stakeholders. They argue that although industry standards and guidelines agree on the importance of LFI, they do not provide enough details on how to this. A result of this is that many organisations struggle with implementing efficient learning systems [73, 29].

An Incident Response Management (IRMA) method is proposed by Line et al. (2008) [55] and further described by Jaatun et al. (2009) [35]. The IRMA method targets integrated operations within the oil and natural gas industry, but is applicable to other industries as well. The framework follows the same basic approach as the standards and guidelines described in Apx. B, but differs in three important aspects:

1. Increased focus on socio-technological interactions between people, processes and technology
2. Increased emphasis on learning from incidents. Both reactive and proactive learning procedures are encouraged.
3. Scoped focus on ICS and SCADA systems within the oil and gas industry

Through interviews with key personnel within the oil and gas industry in Norway, case studies of incidents at petroleum plants on the Norwegian Continent Shelf, and workshops with members of the IRTs of organisations operating within the Norwegian petroleum industry, the team behind the IRMA method developed a thorough understanding of the requirements for incident handling in critical infrastructures.

The IRMA method is structured in three phases as seen in Fig. 2.13:

1. **Prepare**: Plan and prepare for incident response,
2. **Detect and Recover**: Detect indents and restore normal operations
3. **Learn**: Learn from incidents and how they are handled.

Line et al. (2008) and Jaatun et al. (2009) expects organisations to primarily be in the *Prepare* phase, but that incidents will initiate a phase transition to the *Detect and Recover* phase. The *Learn* phase will follow when incidents have been handled.

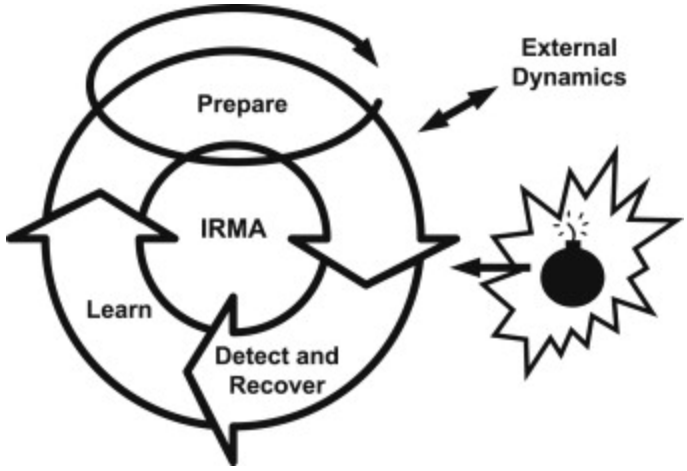

Figure 2.13: The IRMA wheel, taken from Jaatun et al. (2009) [35]. **Prepare**: Plan and prepare for incident response, **Detect and Recover**: Detect indents and restore normal operations and **Learn**: Learn for incidents and how they are handled. An incident is triggered between the Prepare phase and Detect and Recover phase, which is illustrated with a bomb.

To achieve effective detection, recovery and learning operations, Line et al. (2008) and Jaatun et al. (2009) argue that organisations must allocate resources to plan for incidents and conduct proactive learning sessions as part of the prepare phase.

A study by Line and Albrechtsen (2016) further examine the suitability of industrial safety management approaches for information security incident management [56]. They found that doing *lessons learned* after incidents could improve the abilities of societal critical infrastructure operators to foresee future trends and attacks, and that learning which indicators to look for will improve the ability to prevent and detect attacks. They also highlight the need for information to flow beyond IRTs to include larger parts of organisations, and that sharing *lessons learned* would benefit from a systematic approach for learning from incidents.

Grispos et al. (2017) present a solution to enhance feedback and follow-up efforts by integrating lightweight agile retrospectives and meta-retrospectives with traditional reactive incident response procedures [29]. They found that a lightweight agile process *'could be used to drive post-incident meetings in order to collect information that will help answer the queries posed in the NIST guide'* [29]. However, this research does not discuss how *lessons learned* in one organisation could aid a trusted set of peer organisations, nor how organisations could utilise *lessons learned* to identify undetected intrusions.

Lindberg, Hansson and Rollenberg (2010) identified a need for dissemination of *lessons learned*, and consequently added information dissemination as an explicit step in their investigation steps model [54]. This finding was verified by Drupsteen and Guldenmund (2014), who found that sharing of *lessons learned* could trigger a new learning process at the receiving party. Another finding that was verified by Drupsteen and Guldenmund (2014) was that lessons that are learnt by one or more individuals within an organisation could be of significance for the entire organisation [24].

Ahmad, Hadgkiss, and Ruighaver (2012) conducted an exploratory in-depth case study of a global financial institution with the aim of exploring organisational learning in incident response [1]. They found that the lack of an explicit focus on double-loop learning, such as the incident learning system proposed by Cooke et al. (2006) [17] seen in Fig. 2.14, resulted in direct corrective actions rather than fundamental actions that could change the underlying system. As a result, organisations were not leveraging their security experience, nor did their risk assessment processes ingest data about past incidents. Ahmad, Hadgkiss, and Ruighaver (2012) points out that this could be a result of poor communications between related security functions and hence concluded that a double-loop learning model is imperative to learn from past incidents and to avoid repeating intrusions. To achieve double-loop learning, they propose a modified learning model, Fig 2.15, based on the incident learning model by Cooke et al. (2006) [17]. The latter is illustrated in Fig. 2.14. The model proposed by Ahmad, Hadgkiss, and Ruighaver (2012) is based on the measures proposed by Shedden et al. (2010) [73].

Ahmad, Hadgkiss, and Ruighaver (2012) further found that the financial organisation prioritised high-impact incidents rather than high-learning incidents when selecting incidents to investigate [1]. As a result, they propose that both low-impact incidents and precursor incidents should be considered as high-learning incidents. They argue that the risk management team and security strategy & policy developers should be formally included in the incident learning process, both in finding the root cause of incidents and in the information dissemination occurs after incidents. They suggest that, because multiple low-impact incident could be as damaging as a single high-impact incident, clusters of incidents should be handled as a single large incident.

> *'True double-loop learning is only achieved when an organisation is capable of systemically correcting the issues identified through causal analysis. Identifying why a risk has not been identified, or why its mitigation has not been addressed properly, may lead to corrective measures for other risks not directly related to this incident. [...] identifying why a potential incident was not adequately covered by the current security policy, may lead to further improvements to that policy that may prevent future incidents not*

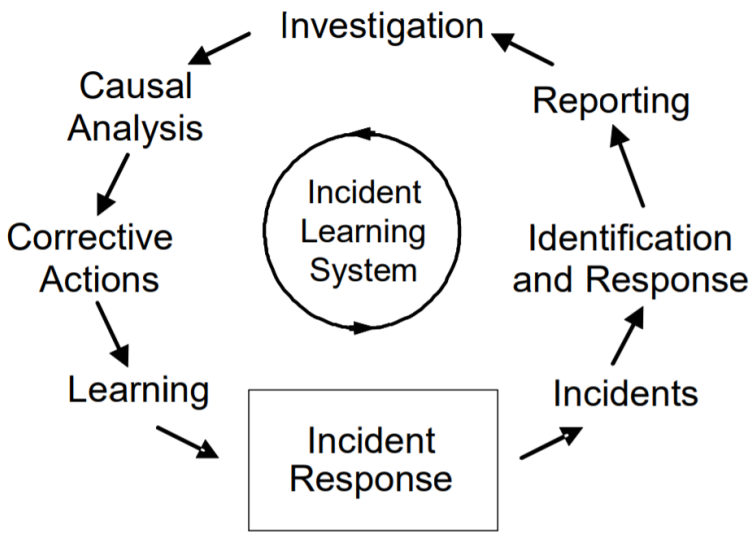

Figure 2.14: The single-loop incident learning system proposed by Cooke et al. (2006) [17], adapted from Ahmad, Hadgkiss, and Ruighaver (2012) [1]. The learning system takes corrective actions to remove unsafe conditions. The corrective actions are taken based on an incident investigation, and the corrective actions are further ingested in the incident learning process. The incident investigation is only used to impose corrective actions, and does not consider the governing variables to the system.

> *directly related to this incident.'*
>
> *(*Ahmad, Hadgkiss, and Ruighaver (2012) [1]*)*

Drupsteen et al. (2013) conducted an empirical survey to identify why organisations find it difficult to learn from incidents [23]. Their study identified that incident reporting and the subsequent evaluation of the incident were the most challenging steps in LFI. Drupsteen and Guldenmund (2014) identified three sub-processes in organisational learning from safety incidents [24]: 1) analysis of events, 2) use of *lessons learned*, and 3) sharing and storing information. The first process should identify the underlying causes for the incident, while the second process should include follow-up steps to leverage the *lessons learned* to improve preventive measures. Their study found that the second process is often neglected and thus reducing the value of the *lessons learned*. The third process involves sharing and storing *lessons learned* as part of the post-incident activities. The literature review that they conducted revealed that there is limited academic literature available on such information dissemination.

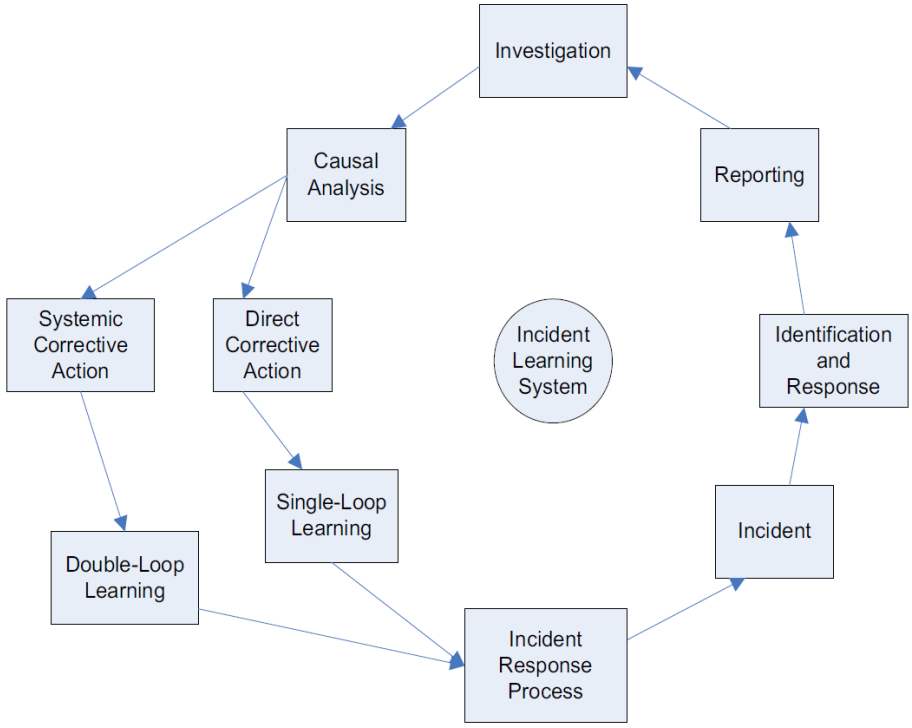

Figure 2.15: A double-loop incident learning system proposed by Ahmad, Hadgkiss, and Ruighaver (2012) [1]. The learning system is a revision of the one described in Fig. 2.14 by Cooke et al. (2006) [17].

In her PhD thesis *Improving organisational safety through better learning from incidents and accidents* [22], Drupsteen (2014) presents a model to use when Learning from Incidents (LFI). The model, seen in Fig. 2.16, illustrates how organisations learn from their own accidents and near-misses. The model is designed to take incidents from other organisations as input as well, thus enabling learning from both in-house and external incidents.

Hove et al. (2014) conducted an empirical case study with three large organisations [30, 31]. They found that the organisations had a difficult time handling information dissemination even though they operated in compliance with industry standards and guidelines. However, their analysis does not take account of information sharing as part of the post-incident activities.

The above researchers argue that organisations do not conduct adequate LFI, either

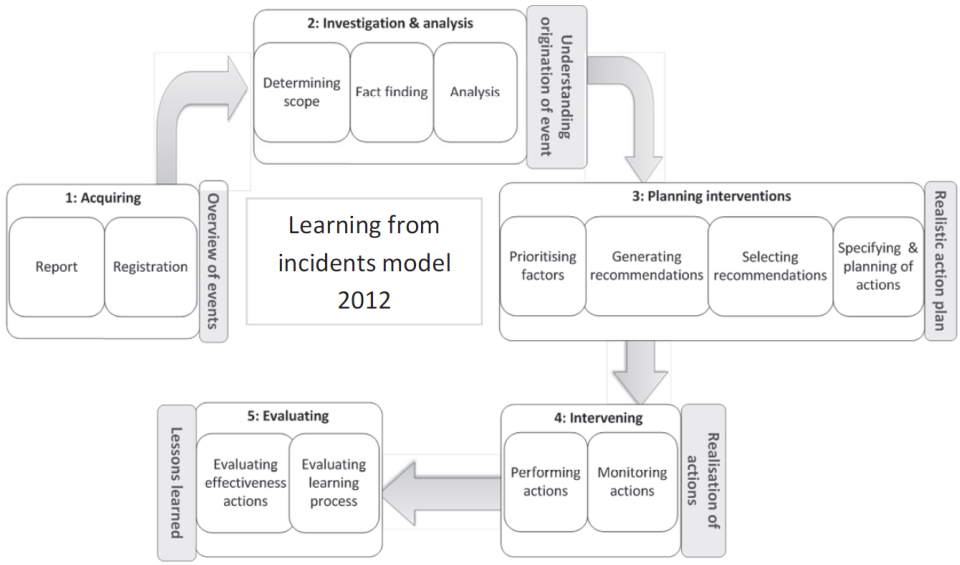

Figure 2.16: Learning from Incidents (LFI) model, taken from Drupsteen (2014) [22]. The first phase, *acquiring*, identifies events from which to learn and handles the registration of the event such that an overview of the events can be fed into the investigation and analysis process. The second phase, *investigation & analysis*, focus on finding the underlying causes of the incident and preventive measures to avoid future recurrence [23], and should give an understanding of the origination of the events. *Planning interventions* involves prioritising the identified contributing factors in the previous phase before generating and selecting recommendations to mitigate the prioritised factors. The prioritisation process could ingest reports of earlier incidents to identify recurrence of contributing factors. The fourth phase, *Intervening*, implements the realistic action plan from phase three. As part of the implementation process, actions are monitored such that they can be adjusted if necessary. The last phase, *evaluating*, evaluates the actions taken and the learning process itself. The end product of this process is *lessons learned*.

due to a lack of allocated resources and effort [55, 35, 1] or because industry best practices and standards lack implementation advice for how to achieve effective learning [73, 29]. However, none of the researchers propose solutions involving sharing *lessons learned* or intelligence generated from intrusions with trusted peer organisations, nor do they address how intrusion data could be used to improve the process of generating threat hunting hypotheses.

Following an in-depth literature review, some key aspects of intelligence-driven incident response and organisational learning have emerged. Firstly, industry standards and guidelines do not provide enough details for organisations to implement efficient learning systems. Secondly, little research on how organisations could exploit intrusions to improve proactive incident discovery are available. Lastly, little research on how organisations could engage in sharing communities, where *lessons learned* are shared both between teams within the organisation and with teams in peer organisations have been conducted.

Chapter 3
Methodology

In the following chapter, the research questions are presented and an overview of the methodology applied in answering these research questions is provided as well as a rational for the choices made. Further, ethical considerations and methodological strengths and weaknesses are discussed.

3.1 Research Questions

The research question for the thesis is:

> How can organisations leverage intrusions to improve their security posture?

The work was guided by dividing the research question into two sub-questions. A solid basis for discussing the main research question was established by answering the following questions:

- How can intrusion analysis help expedite prevention and detection of intrusions?
- How can indicators be used to discover previously undetected intrusions?

Slight changes in the wording of the research question and sub-questions compared to the problem description were done as the thesis evolved. The main research question was changed to include intrusion data in general rather than lessons learned only. Additionally, *root cause analysis* was swapped for *intrusion analysis*, as this is a more accurate term for *security* incidents.

3.2 Choice of Methods

Yin (2009) describe five research strategies and three criteria that can be used when choosing an appropriate research method [81]. The research questions presented in Section 3.1 are 'how'-questions. A descriptive or exploratory study was deemed suitable because neither need to control behavioural events. An exploratory research design was chosen to answer the research questions using an *inductive research approach* due to the innovative nature of the research. Inductive research, or theory-building research, derives theory from observations [6], while *deductive research*, on the other hand, first develops a theory and then evaluate it with observations [61]. Using an inductive research approach made it possible to use observations to derive patterns, rather than evaluate existing hypothesis. A qualitative research method was used to collect in-depth observations from a defined selection of organisations and experts, rather than surveying many organisations and deriving quantitative results. A major advantage of interviewing incident responders and security evangelists is that it gives an accurate representation of how incident response *is* done, and how it *should* be done. We reasoned that it would not be sufficient to survey which standards or guidelines organisations conform to, as this could be inconsistent with their day-to-day processes due to possible mismatches between the governing documents within an organisation and the implemented procedures.

To ensure quality and credibility of the qualitative data collection and analysis, a diversified pool of interview subjects and incidents at two large corporations were studied. Different data collection methods provide cross-data validity checks and capture different perspectives of the same phenomenon [64]. Thus, the data collection methods used to answer the research questions for this thesis were:

- Semi-structured interviews with industry experts
- Case study of the incident handling process and the intrusions handled by participating organisations

Interviews were chosen to get an understanding of how incident response practitioners internationally and across industries are handling intrusions and to gain insight into their experience with using historical intrusions in their work. To complement the findings from the interviews, a case study was chosen to provide insight into practical real-world context of how participating organisations are dealing with intrusions.

The methods are described in the following parts of this chapter.

3.3 Interviews

To answer the research questions, information about how organisations leverage lessons learned after intrusions was gathered. Eight interviews were conducted between January and April 2018. The interviews lasted between 30 minutes and 4 hours each and were all conducted either face-to-face or through video meetings. The interviews aimed at exploring how organisations currently use historical incidents to improve security and what potential the interviewees see in leveraging historical incidents. To achieve this, the interviews were designed to answer how this is currently done in organisations today, what best practices could look like, and what challenges the interviewees would expect for such practices.

The interviews were structured as semi-structured qualitative interviews following a pre-designed interview guide. Thus, the interviews were open and did not follow the interview guide strictly, neither in the sequence of questions asked nor in the relative amount of time spent on each question. This was reasonable since the interviews were one of several information sources used for answering the research questions. The interviewees are, because the interviews were conducted as semi-structured interviews, regarded as participants in the research rather than objects only answering pre-defined questions [30]. The interview guide is provided in Appx. C.

3.3.1 Designing the Interview Guide

An interview guide used in a semi-structure interview defines which topics should be explored during the interview, but does not dictate the sequence of questions nor the amount of time spent on each topic [20, p. 78]. The sequence of questions and the relative time spent on each topic depend on the interviewee's answers and can vary from interview to interview. Nevertheless, the interview guide should be designed to aid the interview to cover all topics during the interview. Depending on the interviewer's preferences, the interview guide could consist of keywords for each topic, complete questions, or a combination of this [77, p. 153]. The interview guide used for the interviews in this project consists of a combination of keywords and complete questions. The questions were used to set the scene and provide context for each topic, while the keywords made it easier to have a natural conversation exploring the topics where the interview subject contributed the most. The interview guide was first developed with full-sentence questions, but most of these questions were reduced to a handful keywords for each topic. This was done to avoid uninspiring and shallow interviews.

3.3.2 Selecting the Participants

The selection of interview participants was governed by what we wanted to achieve with the interviews. From the above discussion, the desire was to get an insight in

the daily operations of incident response, and how the people working on intrusions describe what they believe are best practices in the field. As such, finding participants that are working hands-on with incidents were preferred to get an accurate account of how intrusions are actually handled. Further, finding participants shaping the current field of incident response would give credibility to the proposed best practices of current and future incident response processes. As noted in Sect. 2, Incident Response Teams (IRTs) could be organised in-house or outsourced to vendors specialising in incident handling. It is not obvious whether professionals from these two camps would share the same procedures, perspectives and best practices. For instance, external IRTs being called to handle an intrusion will mostly face successful intrusions because they are called in post-verification of the breach, while internal IRTs, on the other hand, could investigate multiple false-positive intrusions between each true intrusion. To combat this potential perspective bias, the aim was to select participants from a variety of organisations; security vendors, internal IRTs, and national and sector wise IRTs. This selection process is called a 'strategic choice', since the interviewees are selected because of their perceived experience and knowledge [20, p. 74].

3.3.3 Setup for the Interviews

It is easier to pay undivided focus on the interviewee subject during an interview if the interview is recorded [77, p. 166]. Knowing that the interview is recorded can free the interviewer from taking notes and makes it easier to ask follow-up questions rather than trying to remember everything that is being said during the interview. Some of the interviews conducted was audio recorded, while others were not recorded for practical reasons. Interviews being recorded had to be transcribed before analysing their content. This was a time-consuming task. For each hour of recorded audio, the transcription took between eight and ten hours. For the interviews that were not recorded, notes where taken during the interview and elaborated shortly after the interview to get as accurate summaries of these interviews as possible. Additionally, the participants were given these summaries to make sure that nothing was missed or misunderstood from the interviews. None of the participants had any requests for modifications of the summaries.

3.3.4 Presentation of the Interviews

Meaning condensation [41] was used to break down long sentences into shorter and more readable sentences. The findings are presented either as quotes from the respondents or as reproduced summaries of what they said. The presentation of the information collected during the interviews are structured using themes found in the empirical data set, and is similar to the categories used in the interview guide[1]. The approach used for grouping the data is presented in Sect. 3.5.

[1]The interview guide in its entirety is included in Appx. C

The findings from the interviews are presented in Sect. 4.1.

3.4 Descriptive Case Study

Yin (2009) defines a case study as '*... an empirical inquiry that investigates a contemporary phenomenon in depth and within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident*' [81]. Since it would be a focus on best practices in the interviews, a descriptive case study was selected to provide insight in how large organisations handle intrusions and post-incident activities. Sect. 3.4.1 describes the available data sources made available to the author of this thesis in each participating organisation. Document studies and qualitative interviews were used to describe the incident response process in the participating organisations, with a special focus on the post-incident activities.

3.4.1 Introducing the Studied Organisations

In the following, a brief introduction is given to each of the two organisations that participated in the case study. Both organisations have been assigned a pseudonym (Organisation A and Organisation B) to ensure that none of the organisations are identifiable in this thesis. For each organisation, a short break down of the data made available to the author of this thesis is provided.

Organisation A

Organisation A is a large critical infrastructure operator with a global presence. The data made available to the author of this thesis by Organisation A was:

- Informal conversations and interviews with incident handlers, security monitoring personnel, security architects, and senior management governing the security operations within the organisation.
- Documentation of incidents. This included both archived documentation about past incidents and 'live' information about on-going incidents.
- Threat data and information collected about past and on-going incidents.
- Lessons learned presentations and reports from past intrusions.
- Minutes of, and access to, meetings involving reviews of past incidents, and identifying and planning improvements in security operations.

In addition to the data sources listed above, the author of this thesis was given access to sit with Organisation A's Computer Security Incident Response Team (CSIRT)

and Security Operations Center (SOC) from January to June 2018. Such intimate access to the day-to-day operations of these teams provided indispensable insight and value, and facilitated for easy direct observations of how incidents were handled and how the post-incident activities were organised.

To get access to confidential documents at Organisation A, the thesis author had to sign a confidentiality statement. Publication of this thesis required Organisation A's prior written approval, and Organisation A's confidential information should not be published as such. Organisation A had the right to demand reasonable changes, anonymisation of results or omission of Organisation A's confidential information in the publication, for the sole purpose of protecting Organisation A's confidential information.

The findings from the case study at Organisation A are presented in Sect. 4.2.1.

Organisation B

Organisation B is a large Norwegian critical infrastructure operator. The data made available to the author of this thesis by Organisation B was:

- an in-depth incident investigation report. The report is a result of an investigation of the handling of a specific incident. Studying the report gave a thorough insight in the challenges of handling incidents in large organisations and illustrated the means by which the organisation was learning from the incident.

The findings from the case study at Organisation B are presented in Sect. 4.2.2.

3.5 Data Analysis

A *general inductive approach* described by David R. Thomas [76] was used to analyse the collected empirical data. By using this approach, the raw text data is condensed into a brief summary format. This is presented in chapter 4. This approach helps establishing clear links between the research objectives and the summary findings derived from the raw text data. This provides transparency and defendability to these links so that they can be both demonstrated to others and justified given the research objectives. This is fulfilled in chapter 5 were the findings are directly linked with the research questions. Additionally, this approach aids in developing a model about the underlying structures that are found in the data, which is used in developing the model in chapter 6.

The analytic strategies described by Thomas (2006) [76, p. 239-240] guided the interpretation of the findings:

> *Although the findings are influenced by the evaluation objectives or questions outlined by the researcher, the findings arise directly from the analysis of the raw data, not from a priori expectations or models. The evaluation objectives provide a focus or domain of relevance for conducting the analysis, not a set of expectations about specific findings.*
>
> *(*David R. Thomas [76, p. 239]*)*

The transforming of raw data to findings and conclusions were further guided by a process of inductive coding. This process is summarised in three steps [76, p. 241]:

1. Close reading of the text, multiple times.
2. Creation of categories based on identified themes in the data.
3. Revision and refinement of the category system, which involves merging similar categories until there are a handful broad categories.

The final categories from step three can then be used to structure the findings, discussion and conclusion, and provide a basis from which a model or framework is developed.

3.6 Assessment of the Research Methods Used

Research is often assessed on terms of validity, reliability, and generalisability [77, p. 231].

Reliability

Reliability in quantitative research relates to the replicability of the processes and outcome [53]. In other words, the concept of reliability refers to whether other researchers would come to the same conclusions using the same methods [75, p. 202], and whether the research project could be reproduced by another research team [11, p. 250]. In qualitative research, reliability is challenging to achieve using these definitions. Johannessen et. al. (2010) argues that it is inexpedient to achieve reliability in a qualitative research project, as the data collection will be affected by context, such as the experience of the researchers conducting the interviews [36, p. 229]. Leung (2015) further argues that qualitative research should rather aim for consistency, where variations in the results are within some margin of variability [53].

This project used semi-structured interviews to research Learning from Incidents (LFI), and utilised internal reports, direct observations and archival records to complement these findings. For the semi-structured interviews, the interview guide presented in Sect. C was used. If a research team was to replicate this research project using the same interview guide, the results could still differ from the ones provided in this thesis. It is fair to assume that the same interview guide used by a more experienced researcher could guide the interview in another direction, or the interview subjects could provide different answers. This is as expected due to the practical nature of the research and the fact that every organisation is different. To verify the findings from the interviews, the quotes and summaries from each interview were sent to the interviewees. In doing so, the summaries and quotes were verified by the subjects who were given the opportunity to correct misunderstandings.

Because data describing past incidents are regarded as sensitive information by many organisations, collecting large data sets with incident data is difficult. Being limited to one incident investigation report by Organisation B imposed a risk to the reliability and validity of the research, as the data gathered from this report was not verifiable with the means of follow-up interviews or raw data analysis supporting the report. However, since the descriptive case study was used to complement the findings from the interviews with multiple industry experts, the risk was deemed acceptable.

Validity

For qualitative research, validity refers to the appropriateness of the tools, processes and data [53]. This include, among others, whether the research questions are suitable for the desired outcome, if the research design is aligned with the chosen methodology, and that the findings and conclusions are valid for the context and sample. According to Tjora (2017), research is valid when we get answers to the research questions we ask [77, p. 232]. It is paramount that a research project provides transparency in its findings and conclusions in order to assess the validity of the project [75, p. 205]. Tjora (2017) argues that the validity of a research project could be strengthen by providing transparency in the choices of method [77, p. 234], which in this thesis is provided in Sec. 3.2.

Generalisation

An important aspect with research using qualitative research methods is whether the results and conclusions are generalisable and applicable outside the scope of the project. Kvale and Brinkmann (2009) argues that generalisation comes in different flavours [41]. They list three forms of generalisation; naturalistic, statistical, and analytic. With analytic generalisation, the generalisation is dependent on the extent of which the findings in one study can be generalised to another study subject to the same theoretical model [53, 41]. As argued in Sect. 7, the findings in this paper

should be applicable to organisations beyond the data collection scope in this study, subject to the maturity requirements discussed in Sect. 5 and 6 being fulfilled.

3.7 Methodological Strengths and Weaknesses

A benefit of the chosen method is that it gives an accurate representation of what recognised experts within the industry believe organisations should do in order to best learn from incidents. The semi-structured approach to the interviews ensured that the interviews could focus on the specific areas where the subject had expert knowledge. Using a descriptive case study is one of the more practical ways of understanding how incidents are handled in organisations, and challenges associated in doing so.

There are certain problems with the use of interviews and case studies. One of these is that the responses are subjective and thus susceptible to biases. With a small sample size, caution must be applied, as the findings might not be generalisable. Another source of uncertainty is the selection of subjects. It could be argued that the interview subjects that accepted to participate in the study were positive to the need for organisations to learn more from intrusions, and that experts who do not believe in such processes would not find the research interesting.

Additionally, the sensitive nature of documents and archival records describing intrusions complicates data collection during case studies, as exemplified by being restricted to one investigation report in Organisation B. Consequently, the results presented in this study should be interpreted with some caution.

3.8 Ethical Considerations

The main ethical concern related to our research was that confidential information could be revealed during the case study or during the interviews. This could, for instance, be personal data in logs and business infrastructure details in email headers that was part of the collected data for the case study in Organisation A. Further, some information gathered about past incidents during the case study was graded above Traffic Light Protocol (TLP):WHITE[2], and which should not be shared via publicly accessible channels. As mentioned in Sec 3.4.1, the participating organisations were given pseudonyms (Organisation A and Organisation B) and all findings from the case study were anonymised at the end of the study to ensure the organisations cannot be identified.

The researcher is always responsible for obtaining consent from participating interviewees when personal data are collected [75, p. 26]. This was done verbally for each

[2]See https://www.us-cert.gov/tlp for definitions and usage.

participant before their interview. The interview participants that have been named in this report have all given their explicit written consent to be named, and were given the right to withdraw this consent before the thesis was published. The interviewees were offered to read through any relevant sections related to their interview. Further, the project was reported to the Norwegian Centre for Research Data (Norsk Senter for Forskningsdata). All recordings and notes from the interviews were deleted at the end of the study and stored in a secure manner prior to this.

Chapter 4
Findings

In the following chapter, the collected data from the interviews and the case study are presented. The presented findings from the interviews, section 4.1.1-4.1.5, are grouped in five categories: *documentation, intrusion analysis, lessons learned, threat hunting and information sharing*. The grouping is based on categories evolving from the qualitative data analysis, as described in Sect. 3.5. It should be noted that the information introduced in this chapter has not been interpreted or analysed, but presented as it was given in the interviews and the case study.

4.1 Interviews

This section presents findings from the interviews. The interviews have been structured as semi-structured interviews and were transcribed afterwards. The interview participants have diversified backgrounds. Some of them are well-known voices in the global community of proactive defence operations, working in vendor organisations within the information security industry. Others are working as advisers for global incident response teams guiding and supporting organisations in handling information security incidents, or working as incident responders within their respective organisation. The interview participants are:

- **Dr. Adrian Nish.**
  Head of Cyber Threat Intelligence at BAE Systems.
- **Incident Responder 1.**
  Head of Investigations & Incident Response at a global cyber security vendor based in London.
- **Robert M. Lee.**
  Founder and CEO at Dragos Inc, course author of SANS ICS515 – "Active defence and Incident Response" and the co-author of SANS FOR578 – "Cyber Threat Intelligence".

- **David J. Bianco.**
  Incident Detection & Response Specialist. Maintainer of the ThreatHunting Project[1] and a member of the MLSec Project[2].

- **Andreas Sfakianakis.**
  Threat Intelligence and Incident Response professional. Author of the ENISA Threat Landscape (2012) report.

- **Incident Responder 2.**
  Subject Matter Lead of the CERT & SOC at a global operator of critical infrastructure

- **Intelligence Officer 1.**
  Tactical Intelligence Officer at a large Financial Institution

- **Chris Sanders.**
  Founder of Applied Network defence and Author of Practical Packet Analysis & Applied NSM[3]

It is important to bear in mind the possible bias in the responses presented below.

4.1.1 Documentation

> *'All intrusions should be documented. You might even find more value in the failed intrusions than in the successful ones, because by the very nature, those were the first attempts the adversary wanted to do, which means that that might be their go-to Playbook.'* (Robert M. Lee)

Robert M. Lee said he teaches intrusion analysts to use the kill chain model to structure the documentation they make during intrusions. A pitfall for many, he has seen, is that the kill chain is taken too literally, leading to analysts getting confused when there are other steps involved in the intrusion than in the kill chain. He argued that it is important to view the kill chain as a structured schema enabling the analyst to do queries against it for the purpose of intrusion analysis, but that it was never meant to be a documentation of every action of an adversary.

> *'If your expectation is that an incident responder is going to tag every piece of information that they have, that would be ridiculous. However, if you are pulling things off a proxy, you're probably going to be dealing*

[1] http://www.threathunting.net/
[2] http://www.mlsecproject.org/
[3] Network Security Monitoring

> *with Command and Control (C2) traffic. If you are pulling data off some Antivirus (AV) logs, then that's probably exploitation or installation. You can basically pre-tag information [according to the kill chain] based off your data sources and the different security appliances they come in. Then your intrusion analyst just needs to clean up the documentation instead of trying to document everything. If you got to go back post incident and try to document everything, it's not scalable. On the other hand, if you try to pretend that your incident responders and security operations analysts are going to correctly tag all your data, then that's also not a realistic expectation. I usually recommend documenting during the incident, and then cleaning up the documentation post incident.'* (Robert M. Lee)

Incident Responder 1 said that his team most often document intrusions in free-text form during the intrusion, enabling the incident responders to capture whatever information they deem necessary to document. Post-incident they create a summary in the Vocabulary for Event Recording and Incident Sharing (VERIS) format. This is done to detect trends or patterns in the intrusions they handle. Incident Responder 1 noted that trends could arise because of changes in the detection capabilities rather than an actual shift in the threat environment.

Andreas Sfakianakis explained that, in his experience, the best source for intelligence is the ticketing system where all the security incidents are logged. From this source, he said, you can extract information and intelligence based on your requirements and correlate it with external intelligence.

Incident Responder 2 stated that when intrusions are investigated, the investigations have a template for how to document them. Incident Responder 2 was not at liberty to discuss details about the documentation details. He explained that all intrusions that are escalated from their Security Operations Center (SOC) to their CERT are investigated, regardless if the investigation concludes that the intrusion was unsuccessful, and since all investigations follows the same template for documentation, failed intrusions will be documented in the same way as successful intrusions.

> *'An investigation is an investigation. If an unsuccessful intrusion was not investigated, but discarded in the alert triage stage, it will be documented as an unsuccessful attempt during alert triage.'*
>
> (Incident Responder 2)

Andreas Sfakianakis said that the incident responder should collect all the artefacts and documents related to the incident investigation, draft the lessons learned, and close the incident. He further explained that when the incident is closed, or during the

closure of the incident, there should be a handover to the Cyber Threat Intelligence (CTI) analyst who will then take over the follow-up of the incident. Mr. Sfakianakis argued that the CTI analyst should then extract all Indicators of Compromise (IOCs), artefacts, etc., and correlate these with open source intelligence and internal intelligence. This correlation in typically done automatically in a threat intelligence platform.

> *'I structure my report based on the kill chain framework, with a diamond model in each phase of the chain. However, this can often be too much details for the readers, and so then I use a simplified kill chain and free-text to discuss business impact, potential cost of the incident, etc. The report after [big and serious] incidents are at least 50 to 100 pages in my experience.'* (Andreas Sfakianakis)

Intelligence Officer 1 said that their monitoring and detection team escalate intrusions to their Incident Response Team (IRT) depending on the actor, event and target involved. Both failed and successful intrusions are documented if the intrusion is escalated to their IRT. He further explained that intrusions that are not escalated to their IRT are not documented in the same way because it is too resource intensive to document and investigate these intrusions.

4.1.2 Intrusion Analysis

Robert M. Lee argued that every company needs a couple of basic things to best structure their intrusion analysis. Number one, Mr. Lee explained, is that they need a collection management framework.

> *'It doesn't really matter an asset inventory. You should have that, but it's not about knowing your assets. It's about knowing what data you can get from them, what questions those data can answer, and how long you keep that data. You should be able to identify that across your entire organisation. Say for instance you store 60 days of host-based logs in your Demilitarized Zone (DMZ), and your research and development network only have 7 days out of space logs. If you want to ask a question about host-based logs for your R&D network, you only have 7 days even [if the traffic] traverses the DMZ. The value of my intrusion questions are 7 days. Further, say you get some host-based logs from windows; what kind of questions can it answer for me? Well, it can answer exploitation and installation-based activities, but I'm not really going to get a lot of command and control-based activities off of that. Collection management framework is basically your investigations playbook. It is "what can I*

> *do, what are my gaps, what are the questions I can ask?". Then you can identify things like "most of my data is all focused on answering questions about C2 [traffic]." I may not actually have enough investments in other areas to be able to do a real coverage of other parts of the kill chain.'*
>
> (Robert M. Lee)

Mr. Lee further explained that the second thing everybody should have, is a threat model.

> *'The threat model should not be from a vulnerabilities perspective. What I really want to understand is what's most important from the business perspective of my organisation, and what are the threat groups that have shown interest or capability to harm that. From there the focus should be on tradecraft. What tradecraft could impact different portions of my business? Then I map that up to my collection management framework to let me further identify my gaps of the types of questions that I'm going to have to ask for future incidents and for future investigations.'*
>
> (Robert M. Lee)

The third thing that Mr. Lee argued everybody should have, is an intelligence requirements list.

> *'Whether you're going to consume intelligence or generate it, you're going to have intelligence requirements to start that process. What I am actually going to request out of intelligence is going to dictate the type of storing, collection and processing I'm doing with my intrusions.'*
>
> (Robert M. Lee)

In the vein of intrusions, if I can get those three things right Mr. Lee said, I can get really good understanding of every way that I need to structure my intrusion analysis.

Adrian Nish suggested that after intrusions, both major or minors, one should split the kill chain in two: one piece for the phases before the intrusion was stopped and one for the stages after the intrusion was stopped (if any). The first phases, Dr. Nish argued, should help answer how future intrusions with similar tradecraft could be stopped or detected at an earlier phase. He further explained that to achieve this, each phase in the first set of phases should be analysed to identify preventive controls and detection capabilities that would have detected or disrupted the intrusion. Secondly, he explained, if the intrusion failed before the adversary

acted on its target, one should assess what would have happened if the intrusion had not been stopped. Dr. Nish argued that the 'modus operandi' of failed intrusions, or 'hygiene factors', could be used to play the scenario further down the kill chain. This simulation could help detect issues and shortcomings without these being exploited yet.

David J. Bianco explained that the Lockheed Martin Kill Chain [32] is specifically designed to reflect targeted attacks by hands-on-keyboard actors. He explained that the *install* phase is a reference to malware or other enabling tools they might need, but the model was not designed for automated mass-market malware.

> *'Really, there is a mostly-unrealised set of various kill chains for different scenarios. For example, an insider attack would look much different, structurally. That said, it apparently is pretty flexible, because usually what people do is skip steps that don't apply, and anything else that doesn't quite fit that model probably just gets lumped into the "Act on Objectives" phase. The Mandiant Attack Lifecycle*[4] *is also widely used, though less well-known than the Lockheed Martin Kill Chain. The explicit loop structure in the Mandiant model more accurately reflects the lateral movement/recon/compromise cycle the attacker moves through many times while working their way through the network, but it's still mostly based on targeted attacks. Of course, the ATT&CK*[5] *is becoming very popular, though it's still closely aligned with the Lockheed Martin Kill Chain.'* (David J. Bianco)

Robert M. Lee said that during intrusion analysis, you should first get a very granular level overview. Then, he said, you want to get into more details about what specific vulnerability was exploited, etc., but you also want to go one level higher up of it, such that you can get a better return on investment by addressing the root cause. This could for instance be macros in malicious office documents. You could then do further retrospective searches throughout the estate to find similar malicious emails. That, Mr. Lee said, would be threat hunting as part of the process that would now become an intelligence process.

Andreas Sfakianakis accentuated that even though it is challenging and resource intensive to track failed intrusions via AV and Intrusion Detection System (IDS) alarms, this should be done. He reasoned that sophisticated attackers are often seen to be using commodity malware to get into a system if they can, and only use custom tools once they are inside to achieve their objectives.

[4] https://www.fireeye.com/content/dam/fireeye-www/services/pdfs/mandiant-apt1-report.pdf

[5] https://attack.mitre.org/pre-attack/index.php/File:MITRE_preattack_tactics.png

David J. Bianco said that there is value in investigating failed intrusions, but argued that you have to be pretty picky about the things that you would do that extra step.

> *'... you probably can't afford to [investigate failed intrusions] for every potential incident that your team has to investigate. If you have that kind of time you're either super good at security or you're super bad it.'* (David J. Bianco)

> *'[How you review AV and IDS alarms] depends on the size of the organisation and how many of these detections you have. It is different if you have three [alarms] per day or 3000 [alarms] per day. You always have to prioritise, and it depends on how the detection looked like. If it was a detection based on an AV signature, one could export the detection logs from the AV and assess whether it looks like something interesting or not. A cheat sheet*[6] *by Florian Roth can be used to guide this assessment. For instance, if the AV detected and stopped a web shell or Mimikatz, you have to start an investigation by going back in the kill chain and find out how this happened.'* (Andreas Sfakianakis)

Andreas Sfakianakis reasoned that even if you due to a lack of resources cannot investigate an alarm that you would normally prioritise, you should try to identify if there is any way to automate some of the investigation and review the output of the automation regularly.

> *'... if you see that PowershellEmpire*[7] *has been detected by the AV, you first of all need to know how this got past your controls and ended up in one of your workstations. Then you need to go from local to global: you detected something on one workstation, but maybe this is delivered to other workstations as well. Going from local to global is the scoping activity of the incident. Then you need to make sure that no business impact has happened and that everything was blocked during the installation phase. After the remediation phase, one can identify ways of improvement and do the lessons learned session.'* (Andreas Sfakianakis)

> *'During the intrusion you should follow the kill chain both backwards and forwards. Following an incident forwards is also called scoping. Say you have an indicator, [then] you may not be sure at first exactly where on*

[6] https://www.scribd.com/document/346419905/Antivirus-Event-Analysis-CheatSheet-1-1
[7] https://www.powershellempire.com/

> *the kill chain that thing fell, but you found something somehow. Maybe an automated alert, some proactive threat hunting, a tip from the local law enforcement, etc. Not only do you have to work it backwards, but you also usually have to work it forwards to see if there was any following activity that you did not detect. In a lot of cases you will do that first, at least a little bit of that, because you have to know whether you are dealing with just this small piece of activity or if it is an indication of a larger security event that you need to be concerned with.'* (David J. Bianco)

David J. Bianco argued that three security teams within an organisation that could have an interest in simulating the kill chain further based on past intrusions. The first team, he said, are the incident analysts who might want to do such simulations just to improve their own skills, ability to investigate that kind of incidents or knowledge about adversary Tactics, Techniques and Procedures (TTPs):

> *'A lot of times, [past intrusions] have been the basis of the training that we would provide internally. At places that I've worked at previously, we've had our own active internal training and a lot of it would come from intrusions that we've had in the past. Not just the actual events that occurred, but also simulate intrusions for how they could have gone differently. Or, we would take the actual artefacts and show how we did the tear down, and what other capabilities might have been in some of the malware or one of the features of the C2 protocol even though they might not have used in that particular intrusion.'* (David J. Bianco)

Mr. Bianco next pointed out that the CTI team also has a good reason to simulate intrusions when they find a new family of malware that they are particularly interested in:

> *'. . . you'll find threat intelligence teams in the retail sector that try to identify all of the point-of-sale malware that they can, even if it's not targeting the point of sale systems that their retail store uses. It could be something that might not even actually work in your environment, but it might just be something that you need for your general situational awareness. They do this just because they have a vested interest in keeping up with the capabilities of the threat actors, even if they're not targeting their store right now.'* (David J. Bianco)

The third team that Mr. Bianco argued could have an interest in simulating the kill chain based on past intrusions is the endpoint vulnerability management team. He

said that since the endpoint vulnerability management team usually does vulnerability scanning and vulnerability management, they will often want to know how these things work and, if their software has stopped it at the endpoint at this level, what would have happened if that part had not worked on that computer? He explained that the team would have an interest in determining what would have happened if they didn't have a functional AV or the endpoint agent had whitelisted that binary for some reason:

> *'Those three are kind of common, but maybe the endpoint vulnerability management team is somewhat less common than the other two. You will find teams in most mature SOCs that investigate what might have happened, and then use that to turn around and drive improvements.'*
>
> (David J. Bianco)

Mr. Bianco said he would prioritise, and would assume most teams to prioritise, intrusions that were either unusual in some way or very common. The former, he argued, could for instance be a technique that you have not seen before, while the latter could be that you 'just want to dig down and make sure you are getting the most out of it'.

> *'Most of those teams (incident analysts, CTI team and endpoint vulnerability management team) probably have to pick and choose, especially the incident analysts and investigators. The intelligence team and the vulnerability management team might have a bit more time to be more comprehensive, but even then, those guys are going to have to pick and choose what they care about. They're not going to run every sample of the same malware family to the same level.'* (David J. Bianco)

Incident Responder 1 said that he would like to improve the way they do retrospective [intrusion] analysis during or after handling an intrusion, but this is not currently done in a structured manner. He described retrospective analysis to be an analysis of all weak signals in the time-frame of the intrusion. This analysis should be done in the context of what was discovered during the analysis of the main intrusion. The aim of the retrospective analysis should be to uncover paths of the intrusion not yet known, or reveal unrelated intrusions using similar TTPs.

Robert M. Lee reasoned that doing retrospective searches to detect related intrusions can be valuable. He further argued that there are two ways to do a retrospective search; intelligence-driven focus and domain expertise focus.

> *'If you're doing the retrospective search immediately based off what you just learned, it is domain expertise. But if we're going to generate structured hypotheses of similar tradecraft used by the adversary, it has to go to the intelligence folks to do an intrusion analysis first to make sure they codify the right tradecraft, then it comes back as an intelligence generated hypothesis.'* (Robert M. Lee)

David J. Bianco said that what he teaches his analysts to do really early on in alert investigations, is to review a reasonably old set of data to see if they had any previous alerts that they just didn't notice were really malicious or they made a misjudged assessment about. He further reasoned that if you are far enough in the [incident handling] process to have some new knowledge, then you could turn this into a retrospective analysis with new context. However, he pointed out that this could be challenging because you often do this near the beginning of the investigation as part of the scoping to see if you missed anything in the past.

Incident Responder 2 said they always investigate historical [data] based on discovered and extrapolated indicators and TTPs when they have intrusions. Historical intrusions handled by the Computer Emergency Response Team (CERT) are mapped on a TTPs level, while intrusions handled by the SOC are correlated with technical indicators.

Andreas Sfakianakis said that whenever there is a big and serious incident, they do a full analysis of the incident. He further argued that one should go back and do a full scoping of the incident, identifying all compromised hosts, do a forensic investigation, etc.

> *'All failed intrusions related to the successful breach should also be included in this scoping, because these failed intrusions might give you more insight of the adversaries' tradecraft.'* (Andreas Sfakianakis)

Intelligence Officer 1 said that his organisation is, at the moment, only analysing intrusions to improve its prevention, detection and discovery capabilities ad-hoc only due to a lack of resources, but that this is something that they are working on improving.

4.1.3 Lessons Learned

David J. Bianco highlighted that every organisation should have some capability of learning from the incidents that they have, incidents that other people in their environment have, incidents that peers in their industry have, and anything that is

going on in the general internet. This, he explained, would be similar to a threat intelligence capability.

Andreas Sfakianakis reasoned that there should be held formal lesson learned sessions where all main stakeholders are present after every incident.

> *'Every incident should have lessons learned session, but the scale of the session could vary depending on the incident.'* (Andreas Sfakianakis)

Mr. Sfakianakis further explained that findings from the incident are discussed during the lessons learned sessions. These findings, he said, are typically findings targeting technical weaknesses (control gaps) that allowed the incident to happen, but could also be findings about the way the incident was handled (procedures, communication, etc.). Intelligence Officer 1 suggested that a dedicated Information Manager should own the follow-up process, and be responsible for storage and structure of the data generated as part of the process.

David J. Bianco argued that lessons learned, where you review the actions taken during the handling of an intrusion, like what could have done better and what worked well, should be a compulsory step before closing an intrusion as handled:

> *'I think the best organisations don't really consider an intrusion handled until they have had a retrospective discussion about it. [...] in reality, most organisations would say that an intrusion is handled once business operations are restored. '* (Chris Sanders)

Incident Responder 2 said that they do lessons learned throughout the life-span of intrusion investigations. Incident Responder 2 aims for his team to do continuous learning, and he sees this as a step up from doing lessons learned after the intrusion has been handled only. Within his CERT, the goal is to feed lessons learned back to the detection mechanisms. This is done by codifying the lessons learned into detection rules. Incident Responder 2 said that they did not have any formal audits of these detection rules, and that some rules might be present without detecting new intrusions. A reoccurring question during such codification, is whether the organisation has enough data / visibility to act on the new knowledge. Incident Responder 2's organisation has structured this feedback in two loops; one internal loop for quick fixes, and one loop with additional stakeholders for bigger problems. Incident Responder 2 underlined the importance of integrating what has been learned from intrusions into the organisation's risk management, but also stated that technical controls, like modifying firewall rules, could be useful actions based on lessons learned.

Incident Responder 2 further argued for the importance of informing the risk owners about the threat actors that have been involved in intrusions.

> *'We analyse intrusions to understand which IT and business risks we are facing, so that they can be mitigated to the best of our abilities from a cost/value perspective.'* (Incident Responder 2)

Mr. Sanders argued that it's better to hold lessons learned meetings *after* an intrusion has been handled:

> *'I think it's better to do [hold] lessons learned [meetings] after an intrusion has been handled because you should compare the investigation and how it proceeded to what you know [when you hold the meeting]. This way you can ask questions like "why did not discover this until much later?" and "why did I choose to go this route?"'* (Chris Sanders)

Adrian Nish explained how the kill chain could help structure the usage of lessons learned from an intrusion in two ways. First, how could an intrusion have been stopped or detected in an earlier phase of the attack? Dr. Nish argued that each phase of an attack represents an opportunity to prevent or detect the attack. Hence, preventive controls that would have stopped the intrusion and detection capabilities that would have detected the intrusion should be identified for each phase. Secondly, as described in Sect. 4.1.2, Dr. Nish reasoned that you could do lessons learned on what could have happened if the intrusion had not been detected or disrupted.

Robert M. Lee argued that the goal of lessons learned is not only to consume things for active defence[8] or to pass up information to the intelligence people.

> *'The real goal is to push the lesson learned back into the passive defences, and over time, into the architecture. It's almost like a pyramid of how much information you need to be able to push down small amounts of changes. The goal from active defence, or incident response if you will, is to inform intelligence such that they can do their job and inform your own process to become better at your job in active defence. But the real goal is to make as many changes to architectural or passive defences as possible, as long as the changes made to the architecture are built on a really good understanding of risk.'* (Robert M. Lee)

[8] See: The Sliding Scale of Cyber Security [45], Sect. 2.1

Dr. Nish emphasised that the ultimate goal of holding lessons learned sessions is to identify how an intrusion could have been avoided. He stated that some recommended changes are generic and applicable to a range of intrusions; password requirements and policies, configurations, etc. He also said that he values lessons learned sessions that identify and describe the actions taken by the adversary.

Adrian Nish argued that there is a big difference between 'business-as-usual' incidents and full network-intrusion style breaches. He explained that the former are things like routine malware infections, compromised passwords, misconfigurations, brute-force attempts, etc. Adrian Nish further argued that big companies who are 'on-top of their security' deal with this as a 'hygiene factors'; they are not critical but ignoring them can lead to further problems. These may form the bulk of *data* on incidents from some sources, he said, but it is important to separate them from full breaches which actually have business impact.

> *'Big incidents follow an emotional lifecycle. Full network-intrusions where a bad-guy has carried our "hands-on-keyboard" activities, often with Domain Admin level access, can lead to complete loss of control of the network. Even when discovered, the extent of the problem may take weeks or months to understand through forensic analysis, like trying to identify what systems did they access, what data was taken, do they still have backdoors on the network, etc. The management, particularly IT and security teams who feel they have messed up, may go through the full denial, anger, confusion, depression, acceptance, commitment lifecycle during this time. Managing this is as difficult as the technical parts of handling an intrusion, like forensics, malware analysis, etc. "Business-as-usual" rarely looks the same following a big breach. All the "hygiene factor" incidents get investigated, remediated, reported on to management, etc. The team, as long as they are kept motivated, is much better prepared for the next big intrusion if it ever comes.'* (Adrian Nish)

Chris Sanders said that the follow-up actions identified during a lessons learned meeting are often skipped if the severity of the incident is not major enough:

> *'It's most often only the companies that get very largely publicly embarrassed that end up really getting things fixed after intrusions have been handled. It's a shame that it takes bad breaches to get to the point where follow-up activities are actually done.'* (Chris Sanders)

Incident Responder 1 argued that to proactively use incidents to improve security posture, an organisation must have mature security operations in place. This, he

said, include visibility in the network, knowing your inventory and assets[9], having identified your crown jewels, etc.

Mr. Lee said that to him, the structure of the lessons learned process depends on your perspective. He argued that if you are looking at it from a pure incident response perspective, you will usually do lessons learned post-delivery of the intrusion report or post a resolution of the case, but if you are looking at it from an intelligence analyst perspective, you will do feedback at every step and you do it the entire time.

> *'When you are talking about security operations or incident response, you are talking about intelligence consumers. When you are talking about intrusion analysis, you're talking about intelligence production. I usually refer to it as intelligence consumption and intelligence generation. If you are looking at intrusion analysis, you are finding patterns and you are creating intelligence. You are going to do feedback along the way. You are going to do it whenever you learn new things. You might for instance end up finding new tradecraft that you brief your responders on. Feedback or lessons learned is a part of every process when you have enough to articulate. So, it is actually two different schools of thought, depending on whether you are more of an incident responder or an intelligence analyst.'* (Robert M. Lee)

When asked about the life-span of intelligence and information generated from incidents, Robert M. Lee said that it depends on what type of intelligence is generated.

> *'If it is focused on [technical] indicators, I would say it's almost immediately out of date. If we're talking tradecraft, the only time that tradecraft intelligence is not valuable is defined by the adversary. It has an undefined lifetime that I actually think is quite long.'* (Robert M. Lee)

Mr. Lee used an example to explain this further:

> *'We worked on two cases this past year that had two different threat actors. A threat group that we follow broke into an electric utility, got on to the HMIs, took screenshots of the HMIs and exfiltrated them out over DNS. We worked on another case with completely different treat actor, showing different national level interest, that did the exact same*

[9]Incident Responder 1 also noted that GDPR could be related here, as organisations need to have a list of information sources that contains sensitive information. This could improve organisations ability to identify information assets in a shorter timeframe.

> *thing. So, the tradecraft was not specific to the victim and it wasn't even specific to the adversary. To me, when you get to the point of tradecraft, I would articulate that the only time that tradecraft is no longer suitable is when it's no longer possible on the system. We've seen things that have architecturally changed such that the tradecraft is no longer effective, but outside of architectural changes that makes it impossible to do, tradecrafts have the longest life-span of threat information. In this example, HMI screenshot exfiltration is the tradecraft. We can generate detection capabilities for this tradecraft. This tradecraft can be linked with the threat actors that we know have used this, but it is not specific to threat actors.'* (Robert M. Lee)

Mr. Lee further explained the concept of intelligence life-span with the TRITON/TRISIS case.

> *'The TRISIS case was one specific adversary doing something at one very specific chemical plant, where they removed safety logic and what could have been, if the adversary had not messed up, a loss of life situation. A very serious deal. A lot of people have said like "oh TRITON could be repurposed to other victims", but that it's not quite true. Every safety system that's installed, every Safety Instrumented System (SIS), is specific to each and every site, meaning that the attack is not scalable at all to any other sites. The tradecraft of how they modified the logic and how they removed it, however, is infinitely scalable to anybody with a SIS, even on different vendors. Any other adversary could pick that up. So, all indicators related to TRISIS are completely crap and nobody else should use them, but the tradecraft of that case should be considered by everybody in their threat models if they have safety systems.'* (Robert M. Lee)

David J. Bianco reasoned that the best threat intelligence almost always comes from your past intrusions and your security team's direct experience with the adversaries.

> *'In my experience, in many cases the threat intelligence that we used predicted what else the adversary might do, and where else in the environment we should be looking. A lot of this threat intelligence came from our own experience with that threat actor. Sometimes you can get that from a commercial feed, but I think in most cases the commercial threat intelligence providers do not provide the same level of detail.*
>
> *For example, we had an environment that I used to work in that liked to*

> *have multiple copies of various services, even non-critical services as well. If they deployed a print server to service a building, they would deploy two or three identically configured servers. We had one particular threat actor that we noticed always looked for the extra services. If we had an alert show up on one of those services, we always had to look at the ones immediately before and after it as well. If server three alerted, we would typically find them on [server] one and two as well. It didn't happen all the time with that threat actor, but it happened quite a lot and we used that intelligence from our previous cases with them to know that we probably needed to look hard at those related services.'* (David J. Bianco)

> *'... the number one source of intelligence is internal and includes the detections as well as the previous security incidents.'*
>
> (Andreas Sfakianakis)

Andreas Sfakianakis explained that usually the incident responder is the incident response person that is responsible for generating the incident report. He said that the IRT informs the CTI team of what incidents they have worked on, and the CTI team informs the IRT of what threats they have observed and the relevant intelligence they have.

4.1.4 Threat Hunting

> *'Threat hunting is such an ill-defined space. If you ask 10 people how they do threat hunting, you're going to get 10 different answers. There's very little structure to it. [...] even people that are really good at threat hunting are not very good at telling you why or how they're good at it.'*
>
> (Chris Sanders)

Chris Sanders explained that to him, threat hunting breaks down into two categories: exploratory data analysis and TTP based discovery. For the latter, he explained, the hunter is focusing on adversary behaviour and how she can find evidence of such behaviour in her network.

> *'The sources of those TTPs [used for hunting] are going to be widespread. Some of them are going to come from vendors and threat analysis blogs. Another great source, a really great one, are past intrusions. Going through your ticketing system and look for intrusions you can use as input when generating hunting hypotheses is valuable.'* (Chris Sanders)

Mr. Sanders used APT1 as an example. He said that most organisations which have identified APT1 in their networks have done so because FireEye released a report about APT1[10]. The report included the group's TTPs which organisations could use to detect APT1 activity.

> *'Organisations do not necessarily need to have their own dedicated intelligence team to leverage intelligence in threat hunting. Hunters usually have the ability to correlate data from their own network with external intelligence if they can get access to good external intelligence. For me, good external intelligence would be a thorough and detailed report, including a lot of great technical indicators.'* (Chris Sanders)

David J. Bianco argued that threat hunting and incident response have a major overlap. He explained that the main difference is that the threat hunters have a proactive approach that allows them to analyse data to find suspicious indicators that they need to look at. He further explained that the incident investigations, on the other hand, typically start from an alert or external notification, so they don't have to be proactive in the same way or create their own leads.

> *'Once you have a lead, [threat hunters and incident responders] will probably be using the same data and similar tool sets to investigate the lead to find out whether it really was a security incident or not. The hard part is figuring out where the leads are in that giant pile of data, so that is an extra 20% the threat hunters have to do. Otherwise they're very similar and that's one reason, I think, that a lot of successful threat hunters are also successful incident investigators and responders.'*
>
> (David J. Bianco)

Chris Sanders highlighted the two biggest challenges he sees in threat hunting: lack of friendly intelligence and documenting a hunt:

> *'If I had to describe what is the single biggest challenge to most people doing investigative and hunting work it is a lack of friendly intelligence. [...] friendly intelligence it is so important, probably the most important thing in terms of being able to threat hunt successfully. Second biggest challenge in threat hunting is to be able to document a hunt. A hunt is not very different from an intrusion investigation. The input is different;*

[10] https://www.fireeye.com/content/dam/fireeye-www/services/pdfs/mandiant-apt1-report.pdf

> *with intrusion investigation you usually have an alarm or notification telling you where to start. For hunts you have to find out where to start on your own. Once you find something interesting, the mechanics are all the exact same. You ask questions and get answers. You need to have the fortitude and ability to document hunts just as you would with any other investigation.'* (Chris Sanders)

Andreas Sfakianakis explained how threat hunting is different from IOC matching and correlation.

> *'For example, we have automated correlation of indicators. If we receive some indicators, they are automatically correlated. The indicators that we have tagged as interesting, for example Advanced Persistent Threat (APT) indicators, will alert on matches both to the incident response team and the CTI team.'* (Andreas Sfakianakis)

Mr. Sfakianakis said that in addition to automated correlation of indicators, for every report that they get, based on an assessment whether it is relevant, they go through the report and assess how confident they are that they have controls in place for the TTPs mentioned in the report.

> *'For instance, a report could describe a group sending emails with a macro-enabled word document running some PowerShell commands. Then we would ask: If the parent process is word and the child process is PowerShell, can we detect it? Are all controls tuned or not? Then we mainly go to the threat hunting team, but also the incident response team, and then we go "purple teaming", or we go to the data and see if we can detect the activity.'* (Andreas Sfakianakis)

Mr. Sfakianakis described how, in his experience, the CTI team, hunting team, monitoring team and incident responders are interacting.

> *'The work flow is as follows: when we receive an intelligence report, either internally or externally, it is analysed by the CTI team and then the CTI team can give directions or recommendations for hunts to the hunting team. If a hunt turns out to be successful, and we do not expect to have much false positives, the hunt ends up as a monitoring rule. The monitoring rule is picked up by the monitoring/SOC team, which, if it is an incident, escalates the alarm to the incident team. If the incident team*

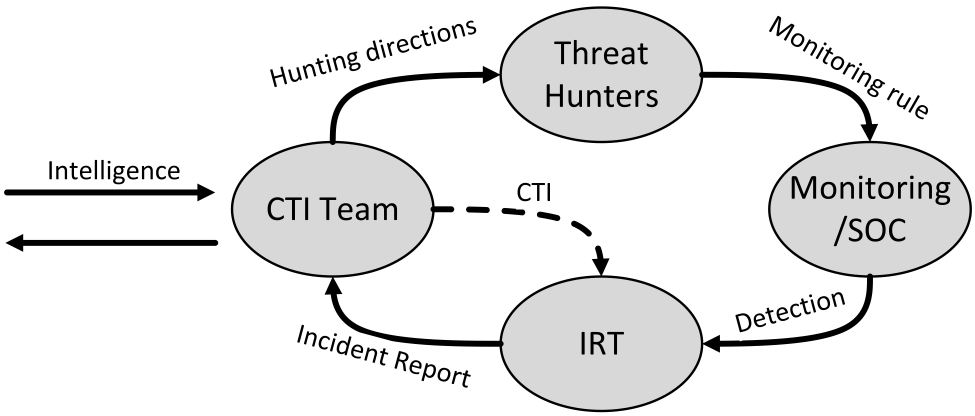

Figure 4.1: The workflow described by Andreas Sfakianakis.

> *assess that it might deal with an APT group, the incident team will engage with a threat management team (CTI team).'* (Andreas Sfakianakis)

The workflow that Mr. Sfakianakis described is visualised in Fig. 4.1.

> *'The output of threat hunting is better detection automation. The ideal thing is [that] you do a hunt once, or maybe a few times, and figure out what actually works repeatedly. Then you work to automate that as much as you can. Maybe it's not possible to make it fully automated, it could be that it's impractical to be automated fully for some reason. For instance, maybe you haven't figured out how to replace at least a little bit of human judgement, or maybe you don't feel like you have the right tools or skillset to close that gap. You might rather look for things that are semi-automated. For instance, instead of admitting an alert, you admit a report that somebody can read and make a judgement call. The more automated you can make it the better, because there's just so many security events going on in the environment all the time that you can't really even hope to keep up by having a person in the loop.'* (David J. Bianco)

> *'I think of the hunt as a cycle. It is not just a thing that you do once. You go through it, you iterate, and you get closer and closer to an automated solution. Eventually, and hopefully, you get there and you can get to the point where you can adjust the true and false positive rates to a ration that you're comfortable with.'* (David J. Bianco)

> *'The automated part is the part that is the most critical for protecting the enterprise. It is the actual functioning system. The non-automated part, on the other hand, is the part that you use to figure out the problems in the automated solutions such that you can automate them better.'*
>
> (David J. Bianco)

Robert M. Lee said he thinks of threat hunting as generating hypotheses and testing them out. Together with David J. Bianco he wrote a paper on generating hypotheses for threat hunting. In the paper, they discuss three ways to do hypothesis generation (See Sect. 2.4.3). Mr. Lee said that if you think about it from the intelligence perspective, you should not be codifying too many lessons learned in it. This is because, he said, your lessons learned should go to intelligence folks and then come back to the threat hunter. Mr. Lee argued that it might benefit you later, but not directly. He went on to explain that the direct aspect [of lessons learned in generating threat hunting hypotheses] is in situational awareness, and even more with domain expertise.

> *'When we go through an incident, we codify lessons learned. There are two ways that this can impact hunting. One is that I now have better domain expertise on my environment, the threats that we are facing, and what I should expect from that. The second thing is that after an incident I should think about it almost in the reverse of hunting, where if I was doing hunting I would have found it before it became an incident. But now that we didn't find it, and it was an incident, what would the hunt look like to be able to identify that incident, and can we put in some automation around the detection of that in the future. So basically, you almost think about it like; my lesson learned is how we could have found this sooner, and can we automate anything at our collection detection processes to make it where this isn't an incident next time.'* (Robert M. Lee)

David J. Bianco argued that you could use past incidents to develop threat hunting hypotheses because domain expertise is based on the analyst's experience in similar situations. He said that most of this experience, either explicit or implicit, is built from past experience with incidents.

> *'If you identified a gap in your security controls as part of an incident investigation, you might want to go and see if any other actors have exploited that gap before it was closed. That does happen, and a lot of times that would be tied to a specific incident where you're trying to find*

> *out if it has other similar occurrences that you might not have noticed before.'* (David J. Bianco)

Mr. Bianco accentuated that it may not always be explicit experience. He reasoned that a lot of times it is based implicitly on your past experiences with incidents, but not with a specific incident in mind.

David J. Bianco said that if they think they have a good handle on who the threat actor [in an intrusion] might be, or even if they think they might have it narrowed down to a few specific threat actors that they have information on, they try to use that information for generating threat hunting hypotheses. Mr. Bianco argued that there are many threat actors that have repeatable TTPs. He explained that it is not necessarily total patterns, but more like 'tendencies', and that if you know those, you should certainly look for them.

The CERT and SOC in Incident Responder 2's company have an operating model with *situational awareness* as one of four main processes, so incorporating knowledge about the organisation is at the core of their thinking according to Incident Responder 2. For instance, they dump their Active Directory database regularly into their SIEM solution so that current and historical privileges, relationships between entities, etc. are easily available for the threat hunters. Although they have additional methods for incorporating knowledge of their organisation, Incident Responder 2 was not at liberty to discuss details about how this is done.

Incident Responder 2 said his team generates threat hunting hypotheses that basically says 'this is how we could be attacked', and that, in his opinion, 'threat hunting is about developing detection, and incidents should always inform detection'.

Incident Responder 1 said that to him, threat hunting is not about using technical threat intelligence to find badness, but rather about looking at a system from an attacker's point-of-view and using knowledge from red teaming, penetration testing and incident response to determine likely paths an attacker might have taken.

4.1.5 Sharing Information

David J. Bianco reasoned that it is probably not of any direct value for an organisation to be sharing information with someone else, but that the value comes when others are giving you something useful. He further explained that if you are sharing with a set of trusted peers in a similar industry, or you have something in common, like a similar technology stack, such that you would expect to have at least some overlap of threat environments, it can be really useful to receive information.

> *'The sharing doesn't have to happen between peers in the same industry, but it could be that they have something else in common that would lead them to have shared adversaries. For instance, it could be something in your supply chain. Or, if you're an energy company, you would have process controllers, and the same process controllers could conceivably be used to brew beer.'* (David J. Bianco)

Mr. Bianco highlighted information about on-going campaigns and information that others are able to derive about the TTPs for some threat actor as the most valuable information to receive.

> *'I used to be in a company that was part of a semi-official organisation for government contractors working for the federal government. It was usually for contractors working for the military, but not always, minus the actual government and military people. A lot of us were working on pieces of the same contracts or with the same technologies, and we had set up this organisation such that we could all have a non-disclosure agreement with that organisation. Such non-disclosure agreement would allow this organisation to be like a clearinghouse for incident information. I would not say we were posting every incident there, but we certainly posted [the] ones that we thought might be of use to our peers in that organisation. Similarly, others would post their incidents there, and we got a lot of good information from this sharing. Not only indicators and detections, but a lot of our threat intelligence that we didn't generate ourselves came from those groups. Some of it was almost as good as if we had generated it ourselves. For example, one of the most common things people would post were samples of phishing emails that they got together with which threat actor they taught sent it. These posts would often contain entire emails together with what they found to be significant indicators for the emails in a phishing campaign. As a first step, we could take those indicators and put them into our detection [mechanisms] if we evaluated that the specific source of that information had done a good analysis. We never blindly took everything that were posted, but rather had to make a decision about the sharing party. A lot of times the sharing party would have been the first people to get hit by that phishing campaign, then we might get hit by it a few days later, or another peer would say that they had seen it a few days earlier, etc. The next step, if any of those initial intrusions were successful, a lot of times they would share information about what the threat actors' next steps were, or they would run the malware and share the capabilities in some malware they found associated with a phishing*

> *campaign. That way, we would get a benefit from their processed or semi-processed intelligence reports as well.'* (David J. Bianco)

Mr. Bianco went on to describe how they could take malware samples [received from peers] and validate their detection capabilities. If they could not detect it, and it got a high enough priority to fix it, then they could initiate a project to improve their detection capabilities to close the detection gap.

Incident Responder 2 said that they are sharing lessons learned with external parties, and that this can include indicators of all levels in the Pyramid of Pain. Incident Responder 2 said that the information that is either shared by them or received from external partners depends on the incident and the adversary, and that in some cases technical indicators could be as good as TTPs.

Incident Responder 2 explained that they share lessons learned with other organisations for reasons such as helping others, building a positive reputation, cultivating a culture of sharing and the hope that they might get something back. When they receive information, tip-offs, like C2 traffic indicating current compromise, is very handy and much valued. He further argued that information on specific threat actors that are relevant for his business are sometimes of value. Incident Responder 2 explained that technical indicators are often not that useful, unless they get them in bulk, but that TTPs are a lot more useful, especially detection mechanisms utilising knowledge of TTPs are highly valued.

Incident Responder 2 argued that for cooperation between external parties to help detect and prevent intrusions, the cooperation has to be between organisations that are at the same maturity level, with the same kind of threats targeting them and with a similar mindset on how to respond to these challenges. A key to information sharing is that the shared information is correct, reliable, timely and relevant.

> *'Cooperation and information sharing in itself is not useful, and can easily do more harm than good if you are up against an APT.'*
> (Incident Responder 2)

Andreas Sfakianakis reasoned that the value of sharing and receiving information from other organisations depends on the size of the CTI team and the maturity of the organisation, and that the appetite of how much an organisation share, how often they do it, etc., depends on this as well.

> *'For most CTI teams, the stakeholders are mainly internal.'*
> (Andreas Sfakianakis)

Adrian Nish said that he believed sharing intelligence was more useful a couple of years ago, but that it is still useful for organisations to receive information about targeted campaigns and new cyber-crime TTPs. He further argued that receiving information early on in a campaign can help the receiving party prioritise which alarms to investigate or do specific hunts if they are provided IOCs. Dr. Nish noted that a challenge with sharing information is that you could end up tip off the attacker and enable the adversary to modify and improve their approach. He said that an evaluation of the consequence should always be conducted before information is shared with external parties.

Chris Sanders argued that information about active campaigns, especially campaigns related to relevant sectors, is the most interesting information for organisations to receive with regards to threat intelligence:

> *'If I'm a CISO of a bank, then I want to know about groups and active campaigns that are actively involved in targeting banks and the financial industry. Campaign data, like who they're targeting, the scope and size of the campaign, what objectives the adversary had where they managed to get a foothold. Basically, anything that can help me build preventive measures and improve detection capabilities for that campaign. I want to know how they're getting in, but I really want to know what they're going for once they are inside and what assets are of most interest to them.'* (Chris Sanders)

Incident Responder 1 argued that sharing knowledge about what is 'normal' in similar environments[11], and information about past incidents, make it easier for the receiving party to detect and understand attacks.

Robert M. Lee said that at his company, they create threat analytics, which is a codification of tradecrafts, and use this in their product.

> *'When we want to share lessons learned to others, that's just an intelligence report. The way to share tradecraft is not in Structured Threat Information eXpression (STIX) and Trusted Automated eXchange of Indicator Information (TAXII). It's not in these formats, but it is an operational level product. With operational level intelligence, generally speaking, you are going to have a diamond model kind of overview, or you going to have some sort of intelligence report. Either way, it doesn't really matter what the format is, as long as the information is there.*

[11] Financial institution networks, Industrial networks, etc.

> *Unfortunately, because of the fanaticism around indicators, all of your models like the STIX and TAXII, as an example, are structured around that and not around the operation level.'* (Robert M. Lee)

> *'I can have a conversation with someone and give them the operational level information they needed. ... the tradecraft of the Domain Name System (DNS) exfiltration of the Human Machine Interfaces (HMIs) give you everything you needed to go find it in your environment. Whether it's a report, a one slide view of a diamond model or a conversation, it's more about the articulation of the tradecraft and the consumption by the analyst who is going to use it, than it is about the formatting of the content.'* (Robert M. Lee)

> *'All CTI teams should have situational awareness, and this is something that can be achieved by consuming intelligence. Specifically, situational awareness on a sector level is important.'* (Andreas Sfakianakis)

Intelligence Officer 1 said that they share incident reports and intelligence reports as part of the dissemination phase of the intelligence cycle. He further argued that it requires a mature organisation to be able to act on intelligence from external parties.

4.2 Descriptive Case Study

So far this chapter has presented findings from interviews. The following section will present findings from the descriptive case study of how organisations handle incidents. Detailed descriptions of how the organisations handled the studied intrusions are not included in this report due to the sensitive nature of these intrusions, and Organisation A- and Organisation B's strict requirements for confidentiality on sensitive information. However, the intrusion resulting in the incident investigation in Organisation B are given a somewhat superficial description. Further, relevant findings from document studies and interviews with members of the IRT and Security Operations Center (SOC) at Organisation A are presented here.

4.2.1 Incidents in Organisation A

The following subsection provides findings from interviews, meetings and document studies of several past and contemporary incidents. Several failed and successful intrusions have been studied as part of this thesis. Although they are only briefly referred to in the following subsection, the discussion in chapter 5 and the proposed model in chapter 6 are built on a thorough understanding of these intrusions. Sensitive

details are omitted from the thesis, and Organisation A has approved publication of the following findings.

The incident response process in Organisation A is a compromise of ENISA's high-level workflow (Sect. B.4), ISO27035 (Sect. B.2), ISO27037 and ISO27043. An incident responder at Organisation A explained that it is difficult to implement a single standard due to how governance within the organisation is structured. The implemented process has the following phases:

- **Detection**
- **Triage** – Deciding what to do, and who will do it.
- **Resolution** – Data Analysis, Resolution Research, Action Proposed, Action Performed, and Eradication & Recovery
- **Closure** – a documentation with a record of the timeline of events, a 'management-readable' summary, and the known root-cause.
- **Post Analysis** – A security review and an IRT review. Changes needed to prevent this from recurring, either technical or procedural, are proposed in addition to reviewing the actions taken by the IRT.

Recommendations from the root-cause analysis should be proposed during the closure-phase if they are deemed urgent, or they will await the post analysis phase. The post analysis phase should take place 1-2 weeks after an incident is closed.

Documentation

Organisation A documents intrusions in free-text format, and categorises the intrusion on the VERIS format both during and after the intrusion has been closed. An incident responder in Organisation A explained that they used the diamond model to document the findings throughout the incident response once, but that it was too resource intensive and time consuming to do on a regular basis. Further, it was explained that this work was not prioritised because a focus on threat actors were sporadic at best. Similarly, Organisation A uses the kill chain sporadically to structure their notes taken during incident response, but that the overhead of doing this so far has not given the expected value. Governing documents state that incidents should be documented in Organisation A, but it is up to each incident responder how they choose to document their analysis.

Failed intrusions are documented on the VERIS format in the same way as successful intrusions, regardless of whether it is escalated to the IRT or not. However, it was explained that the quality of the documentation of intrusions that were not escalated

to IRT was degraded because outsourcing complicated reporting lines and the overall documentation process.

Intrusion Analysis

Retrospective intrusion analysis has been done in Organisation A in the past, but it is not part of the governing procedures and has only been done ad hoc. If IOCs related to an intrusion are made available (by third-parties or trusted peers) before the incident is closed, it is included as part of the scoping and intrusion analysis. If the IOCs are made available after an incident has been closed, there are no procedures or processes in place to assess whether the information is relevant or act on the information. An incident responder argued that a more formalised process should be established if this is to be implemented successfully, and said that technology and processes that automates this would be highly beneficial.

Due to the way security operations are structured in Organisation A, machines infected with malware is often cleaned and/or re-installed before the IRT is able to do a full-scale intrusion analysis. Even though the procedures stated certain conditions for when machines should *not* be re-installed, complex reporting lines and outsourcing complicated this. Thus, information vital for intrusion analysis could be lost. The security team was divided on the reason of this. Some incident responders argued that this was due to a lack of focus to find root causes of malware infections. The SOC team leader explained that the procedures and reporting lines had been updated to facilitate better preservation of data when the root cause of intrusions should be found.

An incident responder at Organisation A explained that they sometimes associate intrusions with specific threat actors. This assessment is often based on their own intrusion analysis with input from external peers and agencies. Several of the studied intrusions in Organisation A were attributed to a set of known APT groups. For instance, a specific threat actor *could* be behind multiple prevented intrusions spanning over several years. A security analyst conducting risk assessments said that they aim to incorporate knowledge about previous threat actors in their risk assessments, but that it is challenging to consume intelligence on threat actors without a structured way of doing so.

Lessons Learned

Organisation A holds lessons learned meetings after each incident, and this is required by the procedures in order to close the incident. However, these meetings are not necessarily held straight after the incident. It could be done weeks or months after the last action was taken. The knowledge generated in the lessons learned meetings

in Organisation A is inserted in a knowledge management system, and is otherwise mainly kept within the participants of the lessons learned meeting.

During the lessons learned meetings, Organisation A mainly focus on:

- The challenges experienced during the handling of the intrusion.
- What worked? What did not work?
- Is the incident ready to be closed?
- Some recommended changes/follow-ups. However, it was noted that these are not necessarily tracked in a structured manner after the incident is closed.

Since IOCs are usually not used after incidents are closed, the challenge of keeping such information up-to-date is not applicable to Organisation A's operations.

Organisation A had an intrusion a couple of years ago, which started with a user opening an attachment from a phishing email. The malicious attachment gathered information about the system it was opened on and documents on the system, like word files and excel spreadsheets. This information was send as encrypted blobs to a cloud storage service. The same cloud storage service was used both for C2 and data exfiltration. Organisation A concluded post-intrusion that only one user had been successfully infected with the malware, but that there *may* have been earlier attempts that failed. Since the same cloud storage service was used for all known C2 traffic and data exfiltration, inspecting the proxy logs was sufficient to conclude that no other *successful* intrusion had occurred. However, because the IOCs that were available for the incident handlers at the time of the intrusion were at the lower levels of the Pyramid of Pain, it was not possible to conclude that all phishing emails that had been sent towards the organisation had been detected. As discussed in Sect. 2.3.1, technical Indicators such as IP addresses, domain names and host- and network artefacts (email subjects and attachment file names), are easy to change. Consequently, related emails could have been delivered without being detected. Thus, an unknown number of phishing emails with malicious attachment *could* have been sent to employees of the organisation without having been detected. Such emails would have been failed intrusions since no C2 traffic was observed.

Threat Hunting

The security analysts in Organisation A doing threat hunting all have long experience from different parts of the organisation. As a result, the analysts have a deep experience with the threat landscape and business mission of the company. Organisation A's SOC team leader explained that the friendly intelligence they have about the

organisation is primarily useful when evaluating the results from hunts, but that they try to incorporate the knowledge when generating hypotheses as well. They mainly use MITRE's Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK) matrix and open source intelligence (OSINT) as inspiration when generating hypotheses. At the moment they do not exchange experiences related to generating hypotheses, but this is something they are in the process of doing by establishing a closed group of potential partners.

Organisation A's SOC team leader explained that they are currently focusing on raising the overall level of detection, but that they within the near future will step up their already existing focus on threat actors that have been involved in past incidents. The focus they have regarding threat actors are TTPs that are known, or assumed, to be associated with the given actor. Currently, much work is done to collect data and establish threat hunting processes and procedures. It was argued that it is a need to mature as an organisation to generate improved threat hunting hypotheses, and that there is an ongoing process in doing this. Enriched and contextualised alarms are often a result from hunts. The SOC team leader explained that alarms and notifications stemming from internal detection systems are enriched and contextualised, and the quality of this metadata is improved as a result of related hunts. A meeting is hosted once a week to share best practices and experiences from past hunts among the members of the SOC. Lessons learned from hunts include identified gaps in telemetry, generated friendly intelligence, improved ways to use tools, and ideas for new hypotheses for future hunts.

Sharing information

Organisation A shares and receives information mainly through personal relations, and to some extent through sector wise sharing forums. Sharing and receiving of information in personal relations happens ad hoc when it is approved by the management in Organisation A. Participation in a sharing forum with trusted peers operating within the same sector takes place in a structure manner, with frequent/weekly group calls and in-person meetings on a regular basis. Shared information is based on a *need-to-know* vs. *need-to-share* tradeoff. The sharing forum was founded six year ago and has developed a deep trust between the members. This trust is paramount for sharing between members in the group. The head of incident response at Organisation A, who has been regularly involved in the activities taking place within the sharing group, said that his organisation had solved cases involving cross-organisation IRT cooperation. Organisation A is participant of a sharing group organised by a national agency, where information and intelligence are shared on a daily basis. Members of Organisation A argued that sharing and receiving information were somewhat dependent on personal relationship and motivation.

Information received by external partners are usually shared internal within the security teams via informal conversations or sporadic emails. Organisation A has presented summaries about past intrusions at conferences. An incident responder in Organisation A said that by participating in sharing forums, he expected to receive 'fresh' IOCs, ideally with context, and additional information that increases awareness within the security teams. Various security teams are using chatting forums and groups to share relevant TTPs, best practices, open source intelligence, etc. internally.

Employees outside Organisation A's security team called for more information about the incidents taking place and descriptions of the organisation's threat landscape. At the moment, the security team provides a dashboard with key metrics aggregating incident data from past intrusions. However, this dashboard was developed for internal use. With senior management requesting access to a dashboard service, it is vital for the security team to provide up-to date metrics and an accurate description past intrusions. The VERIS classification of each incident is the main data source for the current dashboard, and improved quality of the registration is vital for an accurate end-result.

4.2.2 Incident in Organisation B

The following subsection provides findings from a document study of an incident investigation report provided by Organisation B. Sensitive details are omitted from the thesis, and Organisation B has approved publication of the following findings.

Organisation B encountered a breach in their email system in late 2017. The breach led to more than 12,000 emails being exfiltrated to an unauthorised external party. An unknown number of user credentials were also lost to the same adversary. Some of the leaked emails contained confidential information.

Incident description

The first Indicators of Compromise (IOCs) on Organisation B's internal systems are dated at day zero[12]. An internal user account (USER1) was compromised two days later. The user account was used to send phishing emails to a three-digit number of internal and external users. The phishing email contained a link to a fake website asking for user credentials in order to show the user a draft document. USER1 detected unauthorised use of his email account a couple of hours after the phishing emails were sent and reported the incident to Organisation B's IRT. The IRT initiated an incident response process and identified which users had opened the link in the phishing email. All users that opened the link had their password

[12]Details of the timeline are omitted from this report

resat within a couple of hours after the IRT started their analysis, and USER1's user account was disabled for a week before the account was re-opened and the incident closed.

Three weeks later, another user (USER2) received an email that an email he sent could not be forwarded to a gmail-account. USER2 reported the incident to Organisation B's IRT because he had never intended to send the email to a gmail-account, and further verified that he had indeed not sent the email to a gmail-account. The incident was closed by the incident analyst as no malicious traffic could be found on their email server.

A month later, almost 60 days after day zero, it was discovered that seven of Organisation B's user accounts had been configured to forward all emails to the gmail-account mentioned above. A full investigation was initiated 80 days after the incident occurred. The investigation of the incident later revealed that the seven user accounts were compromised on day two of the incident and had been forwarding 12,000 emails before being detected. Because of the time window between the user accounts were compromised and the initiation of the investigation, valuable logs were no longer available for the team doing the investigation. As such, some evidence needed to build the complete picture of the incident were missing.

There are no mentions of lessons learned meetings being held between day zero and day 80 (when the investigation started), but it is fair to assume that if there had been such activities it had been mentioned in the timeline section of the investigation report due to the detailed description of all relevant incident response actions taken in this time window. As such, Organisation B seems to have closed the incident after the recovery steps of incident response (phase 3 in the National Institute of Standards and Technology (NIST) standard) without doing any post-incident activity. Further, the 'early warning' of malicious email activity three weeks after day zero could have been a golden opportunity to limit the damage.

Findings and Recommendations in the Investigation Report

Organisation B's root cause analysis, including their findings into what made the attack possible, is omitted from this report as this information could be valuable information for adversaries.

Organisation B divided the recommended measures into short-term measures and long-term measures. Long-term measures are measures that requires a longer time horizon to implement. The purpose of these measures is that it will require effort over time to establish a good security culture where both technical infrastructure and security operations meet the threat landscape that Organisation B is operating within.

The proposed recommendations are directed at both technical solutions and operative procedures in the organisation. Because the incident was detected at an early stage, but erroneously closed prematurely allowing the incident to evolve, the recommendations did not focus on detection of incidents, but rather preventive technical and operational measures. The process of implementing these recommendations are out-of-scope of this thesis.

Technical short-term measures are focused around architectural changes that would prevent similar incidents to happen in the future. The investigation reports also included operational short-term measures, such as updating password policies and facilitating easy user reporting of phishing attempts.

The long-term measures recommended in the investigation report includes

- performing a vulnerability assessment based on the threat environment that the organisation is operating within
- an audit of the roles and responsibilities between IT, Data & Information management and Security such that an unequivocally perception of who does what, at what time, and why they do it, is achieved
- update strategy documents and associated guidelines for operational security
- improve follow-up and review service level agreements with existing vendors of outsourced security services

It is mentioned in the report that it is not common to include recommended measures to fix problems that were not directly causing the incident, but that an exception is made in this case. It is argued that such recommendations could provide value in the follow-up of the directly related recommendations, and that such recommendations are helpful in demonstrating the complexity in achieving and implementing the proposed recommendations.

Chapter 5
Discussion

In this chapter the findings from chapter 4 are discussed and links between the findings and research questions are established. The chapter starts with a general discussion about the relationship between learning and intelligence, followed by a discussion broken down for each research sub-question. The chapter concludes with a discussion on prominent findings arising from the empirical data.

The following quote by David J. Bianco during his interview nicely highlights the need for organisations to learn from their intrusions:

> *'Every organisation should have some capability of learning from the incidents that they have, incidents that other people in their environment have, incidents that peers in their industry have, and anything that is going on in the general internet. This would be similar to a threat intelligence capability'* (David J. Bianco)

The statement puts emphasis on the need for a learning *capability*, rather than the need for a specific team or tool to achieve it, which is compared to a threat intelligence capability. Intelligence was, indeed, a reoccurring theme in the interviews, with several experts arguing that learning from intrusions should be incorporated in an intelligence generation process. As numerous of the interviewed experts highlighted, intelligence generated from internal intrusions was often found to be among the best intelligence organisations can get. It was argued that because the intelligence is generated internally, the organisation generating the intelligence is free to share it within the organisation without being restricted by receiving conditions. Further, a security team's experience with a reoccurring adversary could be, as demonstrated in the examples provided in Sect. 4.1, highly valuable for incident responders when scoping, collecting data, and in the overall handling of incidents.

Roberts, co-author of Intelligence-Driven Incident Response (2017), emphasise this viewpoint by urging organisations to '*take [their] own incidents, enrich them, understand them and use that as part of [their] own detection, and use that as part of [their] continued enrichment moving forward*' [70].

Intelligence generation *need not*, and maybe should not, be the responsibility of Incident Response Teams (IRTs). However, the security maturity of an organisation varies. Some organisations are on top of their security, with big budgets and well-defined security programs. They might have been through a major breach, like Adrian Nish described in Sect. 4.1.3, and consequently revamped their security posture. Or they might not yet know that they have had a severe breach. Some organisations struggle to keep up with basic security controls. They might not have in-house resources to deal with information security issues, or they lack buy-in from senior management. With this in mind, it should not be expected that organisations with different security maturity would implement such learning capabilities in the same way. A first step could be to make *lessons learned* after intrusions a mandatory step performed by the IRT before closing a case. As the maturity evolves, and the right resources increases, a dedicated intelligence team could be responsible for such capabilities.

5.1 RQ 1: How Can Intrusion Analysis Help Expedite Prevention and Detection of Intrusions?

In order to learn from an intrusion, an organisation need to get a thorough understanding of what happened, why it happened, and how it could have been prevented or detected at an earlier stage. This would require the organisation to possess learning *capabilities*, and, as discussed above, this could be provided by a dedicated intelligence team or by an IRT. The process of analysing an intrusion would thus be part of an *intelligence generation process*.

The intrusion analysis itself could be structured using the Diamond Model of Intrusion Analysis [14] in *combination* with the kill chain proposed by Hutchins et al (2011) [32]. The Diamond Model of Intrusion Analysis would then be used to structure intrusion data within each phase of the kill chain [14, p. 52]. Using the kill chain to structure intrusion analysis, both while responding to intrusions and in the aftermath of intrusions, was indeed a reoccurring theme during the interviews. A common view amongst interviewees was that the kill chain is a good starting point for analysing intrusions with the objective of improving prevention and detection of intrusions. However, some of the experts disagreed on the versatility of the kill chain. While some regarded it as useful only for analysing malware-driven intrusions, others found it useful for a variety of intrusions with hands-on-keyboard attackers. It was argued that the kill chain is flexible if you skip phases not relevant for a given intrusion.

However, the iterative nature of intrusions by Advanced Persistent Threat (APT) actors is not explicitly evident in the kill chain proposed by Hutchins et al. (2011). The attack life-cycle proposed by Mandiant [80, p. 16] takes the iterative nature into account and demonstrates how kill chains could be stacked in order for the adversary to achieve its mission. Regardless of the mean by which the analysis is structured, analysing intrusions to identify improvements to architecture, passive defence and active defence in order to expedite prevention and detection of intrusions would be part of an intelligence *generation* process.

As described in Sect. 2.4.1, intrusion analysis could be done both *backwards* and *forwards* from the point where the intrusion was detected and/or stopped. The interviewed experts agreed on the importance of doing both types of analysis in order to get as much as possible out of each intrusion.

Backwards analysis of the intrusion should be done with emphasis on why the intrusion was not prevented or detected at each phase. Doing a backwards analysis is a key process for finding root causes of an intrusion. The analysis should ideally focus on both the technical and operational reasons for why an intrusion got past a phase without being prevented or detected. For instance, the technical measures to detect Command and Control (C2)-traffic could be in place, but if nobody in the monitoring team is able to look at, or understand, the alarms, the intrusion will still not be detected due to the operational issues.

Forward simulation of the intrusion focus on what would have happened in the remaining phases if the intrusion had not been prevented/stopped. As described by Hutchins et al. (2011), forward simulation enables defenders to prevent future incidents based on Tactics, Techniques and Procedures (TTPs), Indicators of Compromise (IOCs) and capabilities that might not have been evident in the actual intrusion. Imagine a malicious PDF was delivered to a user, and that the Antivirus (AV) engine stopped a dropper from running when the user opened the document. This would have been a failed intrusion, since the preventive measures stopped the adversary from running code on the system. Doing a forward simulation of this intrusion would involve running the dropper in a safe environment, for instance a sandbox, and collect the malware it attempted to download. This malware could then be reverse engineered and analysed. A new capability could be detected, or it could turn out that what looked like commodity malware was actually an attempt by an APT group to gain foothold on the system. This example is technical in nature, but as we will discuss below, it does not need to be only a technical focus in this analysis.

The example in the paragraph above described a *failed* intrusion. Forward simulation

is a great way to learn from failed intrusions[1], or intrusions that did not 'maximise' its potential. Importantly, one should not underestimate the intelligence potential in failed intrusions. The following quote by Robert M. Lee during his interview emphasise this:

> *'You might even find more value in the failed intrusions then in the successful ones, because by the very nature, those were the first attempts the adversary wanted to do'* (Robert M. Lee)

This finding is consistent with that of Ahmad, Hadgkiss, and Ruighaver (2012), who found that low-impact incidents and precursor incidents should be considered as high-learning incidents [1]. This view is further supported by Andreas Sfakianakis, who said that:

> *'All failed intrusions related to the successful breach should also be included in this scoping, because these failed intrusions might give you more insight of the adversaries' tradecraft. [...] you have to be pretty picky about the [failed intrusions] that you would do that extra step. '*
>
> (Andreas Sfakianakis)

Andreas Sfakianakis here mentions an important aspect of analysing failed intrusions; it is resource intensive. The steps described for backwards and forwards intrusion analysis require human resources. Most organisations have too many failed intrusions within any given day to perform a forward simulation of all of them. As David J. Bianco put it; '*[...] you probably can't afford to [investigate failed intrusions] for every potential incident that your team has to investigate. If you have that kind of time you're either super good at security or you're super bad it*'. It is simply too resource consuming to include compulsory forward simulation in incident response procedures. Thus, a priority scheme would be necessary for IRTs to select failed intrusions to put under further scrutiny after they have been deemed to be false positives. David J. Bianco suggested to prioritise intrusions that were either a) *unusual in some way* **or** b) *very common*. The latter, by its very nature, could provide findings that would be applicable to many intrusion attempts. The former, on the other hand, could be applicable for new developments in adversary behaviour.

Forward intrusion simulation is not necessarily only a job for the IRT managing the incident response process, but rather something that multiple teams within an organisation should cooperate on. Threat intelligence teams could collect artefacts from failed intrusions during their collection phase (see the Intelligence Cycle, Sect. 2.3.3)

[1] Failed intrusions are also denoted as *near-misses*

and use forward intrusion simulation as part of their analysis. This would allow them to develop tradecraft and understanding of adversaries targeting their organisation, without waiting for a successful intrusion to occur.

Proper intrusion analysis can find problems that are not directly related to the incident under scrutiny, as illustrated by the investigation report from Organisation B.

In addition to explicit usage of intrusion analysis, it could be used implicitly as training and security awareness material. Onboarding processes could use past intrusion to give new recruits relevant cases to handle. As described in Sect. 4.1, this could be both the actual intrusion, but also simulated intrusions of failed intrusions, similar to forward simulation described above.

5.2 RQ 2: How Can Indicators Be Used to Discover Previously Undetected Intrusions?

One could use Indicators of Compromise (IOCs) on the lower end of the Pyramid of Pain [7] to search the enterprise for intrusions, but the experts concurred with the security manager at Organisation A saying that almost all IOCs they receive are either out-of-date when they receive it or were never relevant for his organisation. This was underscored by Robert M. Lee, who said that:

> *'If it is focused on [technical] indicators, I would say it's almost immediately out of date. If we're talking tradecraft, the only time that tradecraft intelligence is not valuable is defined by the adversary. It has an undefined lifetime that I actually think is quite long'* (Robert M. Lee)

This is in accordance with the principle of the Pyramid of Pain model [7]; technical indicators that are easy to share in data feeds are equally easy for threat actors to change, and thus the IOCs are often only applicable in a single intrusion.

Due to the short timeframe where they are valuable, it is not much value in looking for technical indicators, such as IP addresses or file hashes, to detect future intrusion attempts. However, a set of IOCs could be used to search for undetected intrusions within the time window for where the given set of IOCs are valid. This could be part of threat hunting operations, as discussed in Sect. 2.4.3. With commodity malware, files and C2 servers could likely be re-used, and it is not resource intensive to hunt *lazy* threat actors reusing infrastructure or artefacts by searching for technical indicators. Such operations would be based on the potential for indicator reuse by adversaries leaving the same IOCs in multiple organisations. It is, however, important that

security teams are aware of the limitations of such discovery capabilities. In short, technical IOCs provide little value to prevent or detect future intrusions, but could aid organisations in determining if their network has been compromised by a *known attack*.

According to the Pyramid of Pain [7], TTPs are expected to be more persistent than low-level technical indicators such as file hashes and IP addresses. Receiving TTPs as codified detection logic is highly valuable for IRTs. Such codified detection logic could be run on historical data to discover a yet unknown intrusion within the organisation, or added to the passive defences to expedite prevention and detection of future intrusions. A challenge, though, could be to codify higher-level indicators such that they can be searched for in log data.

Observed adversary tradecraft[2] could be leveraged to discover unknown intrusions since it is located at the top of Pyramid of Pain [7]. Several of the interviewees agreed that tradecraft, and TTPs, are more useful than technical indicators due to it being applicable for a much longer time range and to a wider section of organisations.

Lee (2018) underlines this point by writing '*[...] if you know the tradecraft of the adversary you do not need the indicators to be successful. Indicators aren't inherently bad but how they're used instills false expectations in security and aren't required for defense. [...] If the adversary changes their tools or infrastructure and you scope for indicators they will come back with no alerts. Assuming you aren't compromised at that point would be a bad assumption. Tradecraft is significantly harder to change for adversaries. In short, if someone says to you: "the adversary has expanded their targeting and is using their same methods to go after new sites" a request for indicators is not only not required but can be highly misleading [...]*' [50].

Additionally, it is paramount to distinguish high-quality IOCs from low-quality ones. Several of the experts interviewed argued that vendors tend to choose high-quantity IOCs over high-quality IOCs in the threat (data) feeds they provide. This was backed up by members of the security team in Organisation A, explaining that they had to move carefully in the vendor marked to find threat intelligence services and products that satisfied their requirements.

An organisation could detect an intrusion which is part of an attack campaign, while at the same time peer organisations could be unaware of the same campaign. In such cases, sharing indicators with other organisations could enable them to identify intrusions related to the same attack campaign. It is important not to be tempted into sharing *as many indicators as possible*, but rather keep it at a level where indicators are shared with high confidence and associated context. This is pointed

[2]Tradecraft, defined by Johnston (2005) as '*practiced skill in a trade or art*', often refers to methods and techniques used by adversaries [37].

out by Slowik (2018), who writes '*[. . .] identifying a known-bad hash value or network artifact in historical data allows responders to orient their activity to the perceived threat – so long as the IOC is actually "fleshed out" with amplifying information about the specific threat, its behavior, and other pertinent details. Otherwise, an IOC simply communicates "something bad happened"*' [74].

Finding the root cause(s) of incidents often involve understanding which TTPs were used in the intrusion. Having an understanding of what TTPs are used by adversaries operating in the threat environment as an organisation, could be leveraged when generating hypotheses for threat hunting. As discussed in Sect. 2.4.3, threat hunting hypotheses generally falls into three categories: Intelligence-Driven Hypotheses, Situational Awareness and Domain Expertise. The experts disagreed on whether past intrusions could, or should, be explicitly used when generating threat hunting hypotheses. It was argued that past intrusions should be run through a threat intelligence team before being used to generate hypotheses. Regardless, as a threat hunter, having an understanding of the tradecraft used by her adversaries would improve her situational awareness and domain expertise, thus improving her ability to generate hypotheses.

It was argued that validating the results from a hunt often requires friendly intelligence, or organisational intelligence, about the organisation. Some traffic will be regarded as malicious regardless of which organisation it was found in, but most often the results would have to be assessed with the given organisation in mind. For which users would you expect to find in-bound Remote Desktop Protocol (RDP)-connections? What is normal traffic patterns? These are questions that are easier to answer having friendly intelligence about the organisation.

5.3 RQ: How can organisations leverage intrusions to improve their security posture?

One of the patterns that emerges from the above discussion is that there are mainly two types of learning from an intrusion; the *handling* of the intrusion and the learning about the *threat actor* responsible for the intrusion. The former could be *lessons learned* with regards to root cause of the intrusion, missing data sources, communication lines, resource bottlenecks or telemetry issues. The latter, on the other hand, could be intelligence generated based on the threat actor, and could include behavioural traits or technical IOCs. It could be argued that achieving effective learning of both types would require distinct processes.

The next subsections, therefore, moves on to briefly discuss processes that are required to achieve effective learning of both the handling of an intrusion and the threat actor involved. In addition to learn from intrusions, the outcome of these learning

processes must lead to changes. To account for this, processes required for effective utilisation of the intelligence and *lessons learned* from intrusions are included in the following discussion.

The following five subsections present a brief discussion on lessons learned, intelligence, the need to share, hunting adversary tradecraft and governing variables.

5.3.1 Lessons Learned

Lessons learned sessions in Organisation A were held after an incident were closed. It was, however, challenging for the incident responders to know when to close an incident and hold *lessons learned* sessions. This resulted in an unnecessary long time gap between the last recorded activity related to the incident and the *lessons learned* session, making the *lessons learned* somewhat outdated when they were documented. The outcome of the *lessons learned* sessions were in some cases used to look for signs of compromise of so-far unknown incidents, similar to the procedures described at level 1 of the Hunting Maturity Model in Sect. 2.4.3. However, this was not done according to any method or structural approach and was done only once after the session was held. Surprisingly, the threat hunting procedures described by the Security Operations Center (SOC) team leader in Organisation A indicates that his organisation is at the *procedural* level of maturity, and that they are striving to mature into the more innovative level. One could expect that the organisation would have a structure for searching historical data when IOCs are made available, mapping to level 1 on the scale, given that they seem to be at level 2 in other aspects of their threat hunting program. A reason for this could be that the hypotheses they are developing are not informed by past intrusions, while searching historical data for external IOCs would naturally be guided by past intrusions to filter relevant threat feeds.

Incorporating continuous learning processes within the IRT of Organisation A could ensure that the *lessons learned* from incidents are not outdated once they are identified in a later *lessons learned* session. On the other hand, holding *lessons learned* sessions a while after the incident occurred could, given that the intrusion was part of a larger attack campaign, improve the external intelligence available to the IRT during the session.

5.3.2 Intelligence

Intelligence was a reoccurring theme throughout the interviews and the preceding discussion. Intelligence is described in-depth in Sect. 2.3, and is thus only discussed very briefly in this section. As discussed above, the intelligence process should be directed at uncovering tradecraft, behavioural traits and technical IOCs. As noted by Amann et al. (2012), reliable reporting of complex intrusions today require

external threat intelligence [2]. This, in addition to friendly intelligence as discussed above, should arguably be integrated in an intelligence process consuming incident data to ensure that organisations are learning from intrusions. Such intelligence generation was discussed in Sect. 5.1. The generated intelligence could be used to inform the processes described in the following subsections, namely information sharing, adversary discovery and governing variables.

5.3.3 The Need to Share

Intrusions are not only applicable for learning within the organisation where the intrusion took place. As the interviews and case study revealed, sharing information with trusted peers can be of immense value. Roberts (2018) underlines this point in his talk at the SANS CTI Summit 2018, saying: '*[...] my single favourite thing though, aside from my own incidents, is peer and sharing communities [...] I get so much good intelligence from people that I have gotten to know through events like this*[3] *that face similar problems that I do.*' [70]. This was underlined by David J. Bianco, saying:

> *'[...trusted peers] would post their incidents [in a sharing forum], and we got a lot of good information from this sharing. Not only indicators and detections, but a lot of our threat intelligence that we didn't generate ourselves came from those groups. Some of it was almost as good as if we had generated it ourselves.'* (David J. Bianco)

A key point here is what information is shared, and when it is shared. It was highlighted during some of the interviews that ongoing attacks at peer organisations were helpful in having up-to-date situational awareness and could be used to prioritise and contextualise alarms. Being part of a functioning sharing group could improve the visibility a single organisation has of an attack campaign, going from understanding campaigns from a single point of view to understanding campaigns from multiple perspectives through the eyes of partner organisations. Sharing and improving threat hunting hypotheses could be a convenient way of actionable information without the risk of revealing sensitive information.

As noted by one of the interviewees, receiving tradecraft codified as detection logic is extremely valuable for an organisation to receive, especially if the organisation has observed that specific tradecraft within their networks in the past.

> *'Technical indicators are often not that useful, unless you get them in bulk, [...however,] TTPs are a lot more useful, especially detection mechanisms*

[3] SANS CTI Summit 2018

utilising knowledge of TTPs are highly valued.'

*(*Incident Responder 2*)*

This was shared by Robert M. Lee, who reasoned that one could have a conversation with someone and give them the operational level information they need by sharing observed tradecraft. Such sharing could be a report, a one slide view of a diamond model or an informal conversation, thus not being dependent on a specific format or sharing protocol. This is an interesting view. In a world where automation is regarded as a key to many problems, receiving tradecraft without a well-defined format would be challenging to automate. The focus on behavioural traits rather than technical indicators also accord with our earlier observations, which showed that technical threat data feeds have little value in protecting an organisation.

A recent blog post by Dragos Inc. exemplified how sharing of tradecraft could prove valuable without revealing any sensitive information or details [21]. The blog post, shared widely shortly after being published [10, 66, 63], contained a description of the tradecraft of an APT group known as XENOTIME, and described the groups relation to the TRITON/TRISIS case which was described in Sect. 4.1.3. Notably, the blog post did not provide any technical IOCs, identification of victims or attribution of the group. Regardless, the post provides valuable intelligence that organisations with industrial safety systems could consume to assess and update their threat landscape.

In addition to sharing intelligence generated from intrusions *between* organisations, it could be argued that sharing information and intelligence *within* organisations is valuable. Such sharing could take form as dashboards, showing trends and numbers with regards to intrusions observed by the IRT. Such dashboards could be targeting senior management and decision takers, or they could be used to increase security awareness among employees and contractors. Additionally, sharing could take place in form of in-depth incident reports similar to the one put under scrutiny in Sect. 4.2.2, detailing what happened, how the incident was handled and how it could have been prevented.

5.3.4 Adversary Discovery

'Hunting is a risk because you're betting that there is something there to find – and that you can find it' *(*Caltagirone (2016) [13]*)*

In Sect. 2.3.3 we defined an Advanced Persistent Threat (APT) group to be '*adversaries which have the motivation, resources, and capabilities to sustain malicious effects for a significant length of time against one or more victims while resisting mitigation efforts*' [14, p. 21]. The reoccurring actors that an incident responder at

Organisation A described in Sect. 4.2.1 fit this description, and it would be fair to describe these as APT actors. By definition, these actors will sustain a threat to an organisation over a signification length of time. By leveraging knowledge about threat actors from past intrusions, an organisation could generate threat hunting hypotheses that focus on finding those actors. As discussed in Sect. 5.2, such hypothesis could revolve around IOCs or tradecraft, such as TTPs, associated with a threat actor. This was backed up by David J. Bianco, who during his interview argued that many threat actors have repeatable TTPs that should be used when generating threat hunting hypotheses.

Using IOCs in threat hunting is a disputed topic. Because threat hunting does not have a formal definition of what it means to hunt adversaries, such debates are often based on different perceptions of what it is. In terms of the Hunting Maturity Model [8], described in Sect. 2.4.3, searching historical data with technical IOCs is a very basic form of threat hunting. An organisation could move up in the Hunting Maturity model by hunting for behavioural indicators. Using tools such as the *Cyber Analytics Repository Exploration Tool* (CARET) by MITRE [4], where a hunter can select a set of APT groups and get a mapping of their associated TTPs, could enable organisations to hunt for TTPs by leveraging knowledge about threat actors behind past intrusions.

Adversary discovery nor threat hunting were not mentioned in the investigation report presented in Sect. 4.2.2. Because of the depth of the report, it is likely that this had been mentioned had there been any procedures for this. Threat hunting is not among any of the recommended follow-ups either. This could be due to a need for the organisation to improve their architecture and passive defences before establishing structured threat hunting operations, or because the maturity of the organisation is not yet at a level where active defence is doable. Organisational security maturity is key to establish an active defence program informed by threat intelligence. Lee (2015) argues that '*[. . .] the ability to use threat intelligence requires an organisation be at the point [where] they can use an active defence effectively and accomplish the other components of the [Active Cyber Defense Cycle (ACDC)]. To use an active defence, and in this case specifically Active Cyber Defense Cycle (ACDC), the organisation must first have an understanding and a handle on architecture and passive defences*' [46]. Thus, it could be argued that Organisation B needs to improve their architecture and passive defences before implementing an active defence program. However, the investigation report put under scrutiny in Sect. 4.2.2 indicates a willingness in Organisation B to learn from the incidents. The data provided by Organisation B is not sufficient to determine whether this willingness is a result of a mature security culture within the organisation, or if is is due to a strict safety program requiring investigation reports after major incidents.

[4] https://car.mitre.org/caret/

5.3.5 Governing Variables

Including governing variables in learning procedures is imperative to achieve double-loop learning, as described in Sect. 2.5. It was argued during the interviews that the goal of learning from intrusions should be to improve passive defences on a short term basis, and drive for long term architectural changes. It could be argued that pushing for such architectural changes would require learning feedback into governing variables. The governing variables would be the responsibility of the risk owners in the organisation. Informing the risk owners, and implicitly the governing process, about the threat actors observed in intrusions is vital for them to make informed decisions. Such information would naturally be part of an update of the organisation's threat landscape.

5.4 Summary

It is evident from the findings presented in Sect. 4 and the preceding discussion that the follow-up of an incident should not end with a *lessons learned* meeting as part of the post-incident activities. A structured approach is indeed needed to ensure organisations are able to leverage the data and information generated during an incident response. However, the industry standards and guidelines reviewed as part of a preliminary project contains few recommendations for how to do this [60], a viewpoint backed by several academic papers [73, 29]. A summary of the reviewed industry standards and guidelines is found in Apx. B. It could be argued that the scope of these industry standards and guidelines are mainly the incident response procedures, not the wider security operations taking place in organisations. However, we argue that widening the scope of the industry standards and guidelines to embrace proactive defence principles, such as Learning from Incidents (LFI), intelligence, adversary discovery and information sharing, would aid organisations in structuring their holistic approach to cyber security and make it easier for them to adopt an active defence approach.

The generalisability of these results is subject to certain limitations. For instance, a limited number of experts and organisations were used as data sources. As such, further research should be conducted with increased samples to verify or falsify the findings in this thesis.

Chapter 6
An Extended Incident Response Model

Section 5 revealed that there are currently no models connecting proactive defence with lessons learned after incidents. As such, a model integrating intelligence, adversary discovery and information sharing to the National Institute of Standards and Technology (NIST) Incident Response Life Cycle is therefore proposed in this chapter.

6.1 Intel-Pervaded Incident Response Operations (IPIRO)

The model presented below is an extension of the incident response life cycle presented by NIST [16]. There are no modifications suggested to the activities within the first three phases of the NIST standard. The intention is to enable organisations to implement this extension while at the same time follow the NIST guideline when responding to incidents. The sequential nature of the incident response life cycle proposed by NIST gives the impression that the state of an incident response operation should always follow the arrows between the phases. This is an unfortunate feature of most models; they are simplifications of complex problem. An incident responder should know the model and how the phases relate to each other. However, she should take actions based on the incident in-hand rather than what the next phase in the model is. The same principle applies in the Intel-Pervaded Incident Response Operations (IPIRO) model. The model is a simplification and should not necessarily be followed to the point. It does, however, provide a good starting point for organisations implementing proactive defence and facilitates a shared ground to base future discussion about how organisations could structure Learning from Incidents (LFI).

We have chosen to name the model Intel-Pervaded Incident Response Operations (IPIRO) due to intelligence acting as a glue in the model, interfacing adversary discovery, information sharing with the post-incident activity after an incident response. Each part of the model is explored in the remaining parts of this section.

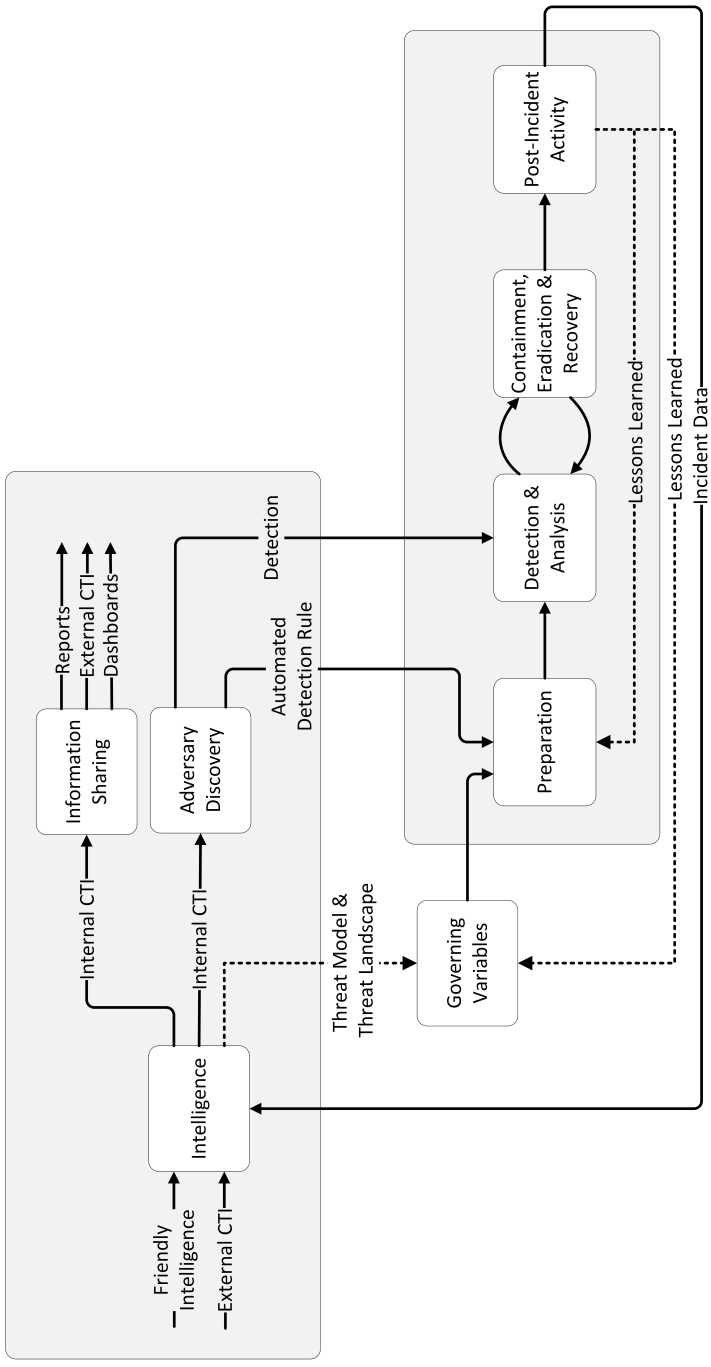

Figure 6.1: Intel-Pervaded Incident Response Operations (IPIRO)

6.1.1 Incident Response

The Incident Response part of IPIRO is ported directly[1] from NIST's Computer Security Incident Handling Guide [16] and which is described in detail in Sect. B.1. The motivation for doing post-incident activities should be to learn from incidents and improve how they are handled. As described in the NIST guideline, the outcome of the post-incident activity should inform and improve the preparation phase of incident response. This is achieved by splitting the output in two categories: lessons learned and incident data. The former, which focus on the *handling* of the incident, is consumed by the governing variables and the preparation phase, while the latter, which focus on the *threat*, is consumed by the intelligence team.

The NIST incident response life cycle has a feedback arrow back to the preparation phase, but the guideline does not, however, include double-loop learning in this feedback. To ensure double-loop learning taking place after incidents, the lessons learned from the post-incident activity is fed back to two distinct elements; the preparation phase and the governing variables for the preparation phase. The **lessons learned** arrows going out from the post-incident activity should focus on how the incident was *handled*, not the adversary involved in the incident. This could for instance be identified bottlenecks in the incident handling that could either be resolved in the preparation phase by introducing a new tool or in the governing variables by redefining how the security teams within the organisation are structured. The latter could be the result of the post-incident activities identifying a need for the security teams to be more proactive or agile in their day-to-day operations.

As discussed in Sect. 5, the best intelligence an organisation can get is from their own incidents. During the handling of an intrusion, the incident responders should make notes about their observations, malware involved in the intrusion, identified Command and Control (C2) traffic and what objectives the adversary *appeared* to have. All this is valuable data to be fed into an intelligence process, but it has to be collected and disseminated from the Incident Response Team (IRT). During the interviews is was argued that incident responders are usually not able to make notes that are of much value to others *during* the incident. Consequently, these notes should be cleaned and elaborated for other members of the organisation to understand them. This should be done by the incident responders themselves during the post-incident activities. To make this process easier, it is advised that the incident responders categorise their notes according to which phase in the kill chain they are investigating[2]. The notes, together with artefacts from the incident, such as spear phishing emails and malware, together forms a bulk of **incident data** that can be processed by an intelligence team. As discussed in Sect. 5.1, incident data should

[1] Except for some minor modifications to the post-incident activity phase.
[2] As argued by Robert M. Lee in his interview, see Sect. 4.1.1

include both successful and failed intrusions. Depending on the amount of successful and failed intrusions an organisation is experiencing, it could be required to prioritise some intrusions to be scrutinised more than others. A specific prioritising scheme is out of scope for this model. The following operations and activities (intelligence, information sharing and threat hunting) should all be run in parallel independent of any intrusions being handled. This might not be obvious from the model illustrated in Fig. 6.1, but is a consequence of the simplifications required for the model to be comprehensible.

6.1.2 Intelligence

The intelligence process of IPIRO embraces the entire intelligence cycle (Ref. Sect. 2.3.3), from direction to feedback. Incident data is collected from, or handed over by, the IRT during or after the post-incident activity. In addition to the incident data, friendly intelligence and external threat intelligence is consumed. This data is processed and analysed as described in Sect. 2.3.3. Friendly intelligence and external threat intelligence put the incident data into context and aid in assessing the organisation's threat landscape and threat model. Intelligence is disseminated to stakeholders during the dissemination step of the intelligence cycle, which in IPIRO are mainly the information sharing, threat hunting processes and the governing variables. Information sharing is described in Sect. 6.1.3 and threat hunting in Sect. 6.1.4. The intelligence generated from an incident could suggest that the threat landscape for an organisation has changed substantially, and thus the governing variables for what is expected of the IRT might need to change. The change in an organisation's threat model could be based on a single incident with a presumed new threat actor, or it could be based on a shift in the intrusions pattern.

Not all organisations should generate intelligence on their own. Some might lack the skills or resources needed, or they might prefer to outsource such operations rather than build an inhouse team to do it. As such, an organisation could adopt the IPIRO model with a third-party handling the intelligence process. However, even if an organisation chooses to do this, it is imperative that the organisation has members of the security team that have intelligence *consumption* capabilities.

6.1.3 Information Sharing

Even though the intelligence process has a dedicated step for intelligence dissemination, it is useful to have a separate explicit step for information sharing. Information sharing consumes intelligence from the intelligence process, but this intelligence might not be prepared to be shared outside the security team and needs to be refined before being shared with a wider audience. It could be argued that this could be handled within the dissemination step of the intelligence cycle in the intelligence process.

However, splitting intelligence generated from incident data and the consumption of the generated intelligence into multiple processes would streamline the overall process should the organisation chose to outsource some of the processes to one or multiple vendors. For instance, as described above, an organisation could choose to outsource the intelligence (generation) process but still handle information sharing outside the security teams. The information that is shared by the information sharing process could have different formats depending on the receiver's preferences and skillset.

Reports based on intrusions and incidents handled within the organisation should be compiled and shared with stakeholders. Stakeholders could be both internal and external ones, including, but not limited to:

- Security Teams, IT Managers, Board Members, C-level executives
- Trusted peer organisations and forums

The reports could either be incident summaries focusing on a single incident, or they could be intelligence reports aggregating information on multiple incidents. The purpose of sharing incident reports with a wider audience is to ensure that the entire organisation has the situational awareness required. It does not imply that everyone should receive the same report, or that everyone has the same need-to-know. However, having a written report about single or aggregated incidents provides an opportunity to inform risk owners, senior management and network operators. The format of the report should be customised to the intended audience. To support audience with different requirements to length and details, the report could be structured with sections targeting executives, management and analysts respectively. However, a report needs not be a three-digit number of pages. It could be a short summary of the incident, what the objectives of the adversary appears to have been, and other relevant information the receiver should know.

External Cyber Threat Intelligence (CTI) could, based on the sharing appetite of the organisation, be internal CTI shared with external parties. The external parties could be the same as the stakeholders for incident reports, but that need not be the case. For most organisations, the main motivation for sharing internal CTI will be the expectation that they will receive relevant CTI in return. The format of the shared CTI could be on machine-readable formats like Vocabulary for Event Recording and Incident Sharing (VERIS) and Structured Threat Information eXpression (STIX)/Trusted Automated eXchange of Indicator Information (TAXII), but organisations should strive to share write-ups of observed adversary behaviour and tradecraft. The latter could, for instance, be a one-slide presentation about what information a threat actor aimed for once inside the network. See Sect. 2.4.2 for a discussion about formats for storing and sharing information. The latter would, for

instance, be applicable for ad-hoc sharing of observed adversary objectives in recent intrusions.

Visualisations represent an opportunity to share aggregated metrics about the incidents handled within the organisation in a way that is easily comprehensible for a wide range of audience. Internal CTI generated by the intelligence process could, for instance, be aggregated on certain adversary attributes such as attack vector and presumed threat actor, but equally valuable could be to aggregate on organisational attributes, such as the targeted business unit (victim). Such visualisations could be used in security awareness projects, be used to detect patterns and trends in intrusions, and is a powerful way of sharing large amounts of data in simple-to-understand diagrams.

6.1.4 Adversary Discovery

The adversary discovery process, renamed[3] from threat hunting that has been used in the preceding chapters, takes internal and external CTI and friendly intelligence as input when generating threat hunting hypothesis. The consumption of intelligence during the adversary discovery process should occur in close collaboration with the intelligence team driving the intelligence process. We will still denote specific discovery operations as *threat hunting* and the analysts conducting these operations for *threat hunters* to align with the literature in chapter 2 even though the overall process has been named *adversary discovery*.

Internal CTI, **External CTI** and **Friendly Intelligence** all contribute with intelligence in intelligence-driven threat hunting hypotheses. Internal CTI is helpful for a threat hunter to know which threat actors are presumed to have been interested in her organisation in the past, which Tactics, Techniques and Procedures (TTPs) have been used and which business units were targeted. External CTI could be used to get updated TTPs on an actor that has not yet been observed in the hunter's organisation, or, in some cases, strong Indicators of Compromise (IOCs) that could be used to find adversaries on the network.

Friendly Intelligence, or organisational intelligence, is intelligence that can aid threat hunters generating relevant hypothesis for discovery and provide good starting points for where to start the hunts. Included in this is the organisation's threat model, crown jewel analysis, information about which elements of an organisation has the greatest risk, and specific operations being undertaken that could change that risk. For instance, a good place to start hunting threat actors could be elements close to the 'crown jewels' in an organisation.

[3] *adversary discovery* is chosen because it distinguish the process from detection and does not inherit marketing 'hype' that could be associated with *threat hunting*.

There are mainly two outputs of the adversary discovery operations: detected intrusions and identified improvements to the current detection mechanisms. Once a threat hunter spots something that *could* be a threat actor, incident response actions could be initiated to verify whether the **presumed intrusion** is in fact an intrusion. Before the IRT is involved, the threat hunter should assess the findings to limit the number of false-positives. However, it could still turn out to be a false positive, in which case the Detection & Analysis phase would terminate the response. If it turns out to be a true-positive, the incident is handled as a regular incident.

The goal of adversary discovery operations is to improve the detection mechanisms by creating new **automated detection rules**. This means that hunters should aim at developing detection logic that could replace future hunts with the same hypothesis.

6.2 Responsibilities – Who Does What?

IPIRO could be implemented as a shared initiative between the security teams within an organisation. This would include, but not be limited to, Security Operations Center (SOC), Incident Response Team (IRT), intelligence team and risk team. Smaller organisation might not have dedicated personnel for each of these teams, and others will be outsourced. Hence, the discussion about who does what in IPIRO in terms of teams depends on the organisation, and how it has structured its security teams. Thus, instead of thinking about the different teams required to implement IPIRO effectively, it makes sense to think about the *capabilities* an organisation should have. In the following discussion, a *team* does not need be a formalised group within the organisation but could be a group of personnel that have a shared responsibility for some task. An organisation could have *one* security team but implement multiple capabilities within that team.

All organisations should have incident response capabilities. This could be implemented as an in-house Incident Response Team (IRT), or it could be bought as a service from a third party. Being the glue in the model, having some intelligence capability is paramount for an organisation to successfully implement the principles of IPIRO. As with incident response, this need not be in-house, but could be outsourced. However, as discussed in Sect. 6.1.2, even if the intelligence capabilities are provided by a third party, some intelligence consumption capability should be in-house. Adversary discovery could be done by a dedicated threat hunting team, or it could be a collaboration between the Security Operations Center (SOC) and IRT. As pointed out by Lee and Lee (2016), hunters provide most value to an organisation when they are allowed to be dedicated at finding threats rather than having to fix network configurations they might discover as part of a hunt, or respond to monitoring alarms during hunts [52, p. 7]. Consequently, if an organisation implementing IPIRO makes adversary discovery a collaboration between its SOC and IRT, awareness of the need

for dedicated hunters are required. As with adversary discovery, the information sharing could be done either by a dedicated team or by the IRT.

6.3 Rationale

'All models are wrong, but some are useful' (George E. P. Box)

The above quote is applicable to IPIRO and could be further extended to 'all models are wrong, but some are useful*, sometimes*'. The model is not a blueprint for the security operations of an organisation. The relationship lines connecting the phases in the model are *not strict*, nor are they all-encompassing. The IRT will receive intelligence from the intelligence process during an incident, although this is not explicitly illustrated in Fig. 6.1. This does not imply that the IRT seldom find intelligence useful during incident response. Organisations could implement the principles of the model, but analysts implementing the model should always know *why* an activity is done other than that it is listed in the model. A model of a complex process must be simplified in other to be useful. Relationship lines could have been draw between all phases and processes in the model, but this would have made the model useless. A reason for this is found the literature. Jaatun et al. (2009) and Cusick and Ma (2010) found that having simple plans for incident management was preferred to comprehensive and complete plans [35, 19, 31]. Cusick and Ma (2010) argued that simple plans could easily be explained and implemented, and that systematic plans could be leveraged to collect metrics about incidents [19, 31]. From these findings it is evident that a simplified model is more useful than a comprehensive model taking edge-cases into account.

As discussed in Sect. 5, organisations should make sure that they are able to do basic security operations before they implement a model such IPIRO. Describing specific requirements for *basic security operations* is out of scope of this thesis, but should, at the very least, include network and asset visibility, and having a threat model, a list of high-value assets and defined reporting lines.

A drawback with the model, as discussed in Sect. 6, is that it supports learning 'after-the-fact', rather than explicitly through the life cycle of an incident. A natural progression of this work is to analyse if and how continuous feedback from the incident handling could be incorporated in the model. In spite of its limitations, the model certainly adds a set of new features to the current incident response standards. Implementing and evaluating the IPIRO model in one or multiple organisations would put the model under further scrutiny, and would be necessary to explore the real-world implications of the model.

Chapter 7
Conclusion

The objective of this study was to investigate how organisations can leverage intrusions to improve their security posture. By interviewing leading experts in the incident response and threat intelligence community, a thorough understanding of experts' recommendations and experience were gained. Further insight was gained by studying an investigation report on a high-impact incident in a critical infrastructure operator. Observing and interviewing members of the security teams of a global operator of critical infrastructure provided an appreciation of the challenges associated with handling incidents in complex environments and the difficulties in adhering to industry standards and guidelines on incident response.

By analysing the findings, this study offers new insight into how organisations can learn from intrusions. The study indicates that intelligence generated from internal intrusions are often the best intelligence an organisation can get. Organisations should strive to exploit internal intrusions, and intelligence derived from intrusion analysis, with a structured approach to continuously expedite prevention and detection of intrusions. The research has also shown that there is a lack of structured methods for organisations to leverage lessons learned after intrusions. In particular, there is a lack of methods integrating double-loop learning with proactive discovery or information sharing with internal and external stakeholders. There are, nonetheless, models that structure *either* organisational learning *or* intelligence-driven active defence. One consequence is that high-value intelligence generated from the intrusion data is not used effectively, or not used at all, when generating threat hunting hypothesises. Further, without a structured approach for sharing intelligence, stakeholders that could have acted on that intelligence are instead making less informed decisions.

We argue that to achieve effective learning from intrusions, intrusions must be documented and shared in a technical and non-technical context. The present study establishes a model for solving the deficiency of structured ways for organisations to leverage intrusions. The model should be regarded as guidance for organisations implementing or structuring Learning from Incidents (LFI), or combining this with

threat hunting and information sharing. To combine organisational learning with intelligence-driven active defence, the proposed model distinguishes lessons learned about the handling of an intrusion from lessons learned about the threat actor involved. The former, a handling review, is fed into both governing variables and security preparations to achieve double-loop learning. The latter, a threat review, is fed into an intelligence process were intelligence is generated based on incident data provided by the Incident Response Team (IRT). The intelligence process, which collected incident data, friendly intelligence and external Cyber Threat Intelligence (CTI), provides input for three distinct processes; information sharing, adversary discovery and governing variables. Information sharing is a dedicated process for sharing information based on the intelligence generated in the intelligence process. The sharing could be of various formats, with both internal and external stakeholders. Adversary discovery consumes intelligence from the intelligence process to generate hypotheses for threat hunting. The adversary discovery process aims at developing automated detection logic for hypotheses where applicable and if a hunt leads to a finding of malicious activity, an incident response process should be initiated. The last process, creating governing variables, consumes threat landscape updates from the intelligence process and lessons learned about the handling of intrusions. Based on this feedback, changes to the variables governing the preparation phase could be made. As discussed in chapter 6, the purpose of using an extended incident handling model, like the Intel-Pervaded Incident Response Operations (IPIRO) model, is to better utilise the output of the post-incident activities.

Further work is required to measure the impact of the IPIRO model as compared to the original National Institute of Standards and Technology (NIST) Incident Response Life Cycle. There is room for further improvements with regards to determining organisational requirements before implementing the model. It could be that organisational structures could affect the way an organisation learn, or how the organisation should prioritise failed intrusions for a wider intrusion analysis. Distributed organisations might need to take a different approach than centralised ones. Requirements for the intelligence flowing out from the post-incident activity phase in the IPIRO model should, if found appropriate, be formalised and further described. If the subject field is to be moved forward, a better understanding of how intrusion analysis should be structured after intrusions to best facilitate learning have to be developed. Further studies could assess how proposed variants of the kill chain and attack life cycle models are applicable to different types of intrusions, ranging from commodity malware intrusions to sophisticated nation-state adversaries. Results from such research could be helpful for organisations in choosing how to structure intrusion analysis during the post-incident activity and move the research area forward.

References

[1] Atif Ahmad, Justin Hadgkiss, and Anthonie B Ruighaver. Incident response teams–challenges in supporting the organisational security function. *Computers & Security*, 31(5):643–652, 2012.

[2] Bernhard Amann, Robin Sommer, Aashish Sharma, and Seth Hall. A lone wolf no more: supporting network intrusion detection with real-time intelligence. In *International Workshop on Recent Advances in Intrusion Detection*, pages 314–333. Springer, 2012.

[3] Chris Argyris and Donald A. Schön. *Theory in practice : increasing professional effectiveness*. Jossey-Bass, 1974.

[4] Sean Barnum. Standardising Cyber Threat Intelligence Information with the Structured Threat Information eXpression (STIX). Technical report, MITRE, February 2014. Available at: https://www.standardscoordination.org/sites/default/files/docs/STIX_Whitepaper_v1.1.pdf Accessed: 22.05.2017.

[5] Irma Becerra-Fernandez, Avelino J González, and Rajiv Sabherwal. *Knowledge management*. Pearson/Prentice Hall, 2004.

[6] Anol Bhattacherjee. *Social science research: Principles, methods, and practices*. Global Text Project, 2012.

[7] David Bianco. The Pyramid of Pain, March 2014. Available at: http://detect-respond.blogspot.no/2013/03/the-pyramid-of-pain.html Accessed: 27.03.2018.

[8] David Bianco. A Simple Hunting Maturity Model, October 2015. Available at: http://detect-respond.blogspot.no/2015/10/a-simple-hunting-maturity-model.html Accessed: 27.04.2018.

[9] David J. Bianco, 2016. Tweet. Available at: https://twitter.com/RobertMLee/status/996559611133071361 Accessed: 31.05.2018.

[10] Chris Bing. Trisis masterminds have expanded operations to target U.S. industrial firms, May 2018. Available at: https://www.cyberscoop.com/xenotime-ics-cyber-attacks-trisis-dragos/ Accessed: 25.05.2018.

[11] Svend Brinkmann and Steinar Kvale. Det kvalitative forskningsintervju. *Gyldendal Akademisk*, 2009.

[12] Scott Buckler and Nicholas Walliman. *Your dissertation in education*. Sage, 2016.

[13] Sergio Caltagirone. Building Threat Hunting Strategies with the Diamond Model, October 2016. Available at: http://www.activeresponse.org/building-threat-hunting-strategy-with-the-diamond-model/ Accessed: 27.05.2018.

[14] Sergio Caltagirone, Andrew Pendergast, and Christopher Betz. The diamond model of intrusion analysis. Technical report, Center for Cyber Intelligence Analysis and Threat Research Hanover MD, 2013.

[15] D Chismon and M Ruks. Threat intelligence: Collecting, analysing, evaluating. *MWR InfoSecurity Ltd*, 2015.

[16] Paul Cichonski, Tom Millar, Tim Grance, and Karen Scarfone. Computer security incident handling guide. *NIST Special Publication*, 800(61):1–147, 2012.

[17] David Cooke, Robert Lee, Sandra Iftody, Erin McKimmon, Jody Powers, Peter Dunscombe, Meina Dubetz, and Rahim Heshmati. A reference guide for learning from incidents in radiation treatment, 2006. Available at: http://www.assembly.ab.ca/lao/library/egovdocs/2006/alhfm/153508.pdf Accessed 06.05.2018.

[18] Jason Creasey. Cyber security incident response guide, 2013. Available at: https://www.crest-approved.org/wp-content/uploads/2014/11/CSIR-Procurement-Guide.pdf Accessed: 05.05.2018.

[19] James J Cusick and Gary Ma. Creating an itil inspired incident management approach: Roots, response, and results. In *Network Operations and Management Symposium Workshops (NOMS Wksps), 2010 IEEE/IFIP*, pages 142–148. IEEE, 2010.

[20] Olav Dalland. *Metode og oppgaveskriving for studenter*. Gyldendal akademisk, 2017.

[21] Dragos. XENOTIME, May 2018. Available at: https://dragos.com/blog/20180524Xenotime.html Accessed: 29.05.2018.

[22] Linda Drupsteen. *Improving organisational safety through better learning from incidents and accidents*. Copenhagen: Centre for Industiral Production; TNO, 2014.

[23] Linda Drupsteen, Jop Groeneweg, and Gerard IJM Zwetsloot. Critical steps in learning from incidents: using learning potential in the process from reporting an incident to accident prevention. *International journal of occupational safety and ergonomics*, 19(1):63–77, 2013.

[24] Linda Drupsteen and Frank W Guldenmund. What is learning? a review of the safety literature to define learning from incidents, accidents and disasters. *Journal of Contingencies and Crisis Management*, 22(2):81–96, 2014.

[25] Robert Duncan. Organizational learning: Implications for organizational design. *Research in organizational behavior*, 1:75–123, 1979.

[26] EY. Cybersecurity regained: preparing to face cyber attacks, 2018. Available at: https://www.ey.com/Publication/vwLUAssets/ey-cybersecurity-regained-preparing-to-face-cyber-attacks/\protect\T1\textdollarFILE/ey-cybersecurity-regained-preparing-to-face-cyber-attacks.pdf Accessed: 22.05.2017.

[27] Gartner. Reviews for Security Threat Intelligence Products and Services. Available at: https://www.gartner.com/reviews/market/security-threat-intelligence-services Accessed: 22.05.2017.

[28] Jimmy A. Gomez. The Targeting Process: D3A and F3EAD, 2011. Available at: http://www.dtic.mil/dtic/tr/fulltext/u2/a547092.pdf Accessed: 04.05.2018.

[29] George Grispos, William Bradley Glisson, and Tim Storer. Enhancing security incident response follow-up efforts with lightweight agile retrospectives. *Digital Investigation*, 22:62–73, 2017. Available at:https://www.sciencedirect.com/science/article/pii/S1742287616301293 Accessed: 22.04.2018.

[30] Cathrine Hove and Marte Tårnes. Information security incident management: an empirical study of current practice. Master's thesis, Institutt for telematikk, 2013.

[31] Cathrine Hove, Marte Tarnes, Maria B Line, and Karin Bernsmed. Information security incident management: identified practice in large organizations. In *IT Security Incident Management & IT Forensics (IMF), 2014 Eighth International Conference on*, pages 27–46. IEEE, 2014.

[32] 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. *Leading Issues in Information Warfare & Security Research*, 1(1):80, 2011. Available at: https://www.lockheedmartin.com/content/dam/lockheed-martin/rms/documents/cyber/LM-White-Paper-Intel-Driven-Defense.pdf Accessed: 22.05.2018.

[33] Information Security Forum. You could be next - learning from incidents to improve resilience, 2012.

[34] International Organization for Standardization. ISO/IEC 27035-1:2016(E) - Information technology - Security techniques - Information security incident management - Part 1: Principles of incident management - First edition, 2016.

[35] Martin Gilje Jaatun, Eirik Albrechtsen, Maria B Line, Inger Anne Tøndel, and Odd Helge Longva. A framework for incident response management in the petroleum industry. *International Journal of Critical Infrastructure Protection*, 2(1-2):26–37, 2009.

[36] Asbjørn Johannessen, Per Arne Tufte, and Line Christoffersen. *Introduksjon til samfunnsvitenskapelig metode*. Abstrakt Oslo, 2010.

[37] Rob Johnston. Analytic culture in the US intelligence community: An ethnographic study. Technical report, US Central Intelligence Agency, Center for Study of Intelligence, 2005.

[38] *Joint Publication 3-13: Information Operations*. United States. Joint Chiefs of Staff, February 2006. Available at: https://www.hsdl.org/?view&did=461648 Accessed: 26.04.2018.

[39] Charles Schmidt Julie Connolly, Mark Davidson. The Trusted Automated eXchange of Indicator Information (TAXII). Technical report, MITRE, February 2014. Available at: http://taxii.mitre.org/about/documents/Introduction_to_TAXII_White_Paper_May_2014.pdf Accessed: 22.05.2017.

[40] Patrick Kral. Incident handler's handbook. Technical report, SANS Institute Information Security Reading Room, December 2011. Available at: https://www.sans.org/reading-room/whitepapers/incident/incident-handlers-handbook-33901 Accessed: 31.10.2017.

[41] Steinar Kvale and Svend Brinkmann. Interviews: Learning the craft of qualitative research. *California, US: SAGE*, pages 230–43, 2009.

[42] Robert M. Lee. Active Cyber Defense Cycle. BSides Huntsville Alabama. Available at: https://youtu.be/MkuH7FJOO_s Accessed: 13.04.2018, 2015.

[43] Robert M. Lee. Active cyber defense cycle: Asset identification and network security monitoring, 2015. Available at: https://www.plantengineering.com/single-article/active-cyber-defense-cycle-asset-identification-and-network-security-monitoring Accessed: 27.04.2018.

[44] Robert M. Lee. Data, Information, and Intelligence: Your Threat Feed is Not Threat Intelligence, 2015. Available at: http://www.robertmlee.org/data-information-and-intelligence-your-threat-feed-is-not-threat-intelligence/ Accessed: 27.04.2018.

[45] Robert M. Lee. The sliding scale of cyber security. White paper, SANS Institute Information Security Reading Room, August 2015. Available at: https://www.sans.org/reading-room/whitepapers/analyst/sliding-scale-cyber-security-36240 Accessed: 26.03.2018.

[46] Robert M. Lee. Threat Intelligence in an Active Cyber Defense (Part 1), 2015. Available at: https://www.recordedfuture.com/active-cyber-defense-part-1/ Accessed: 27.04.2018.

[47] Robert M. Lee. Threat Intelligence in an Active Cyber Defense (Part 2), 2015. Available at: https://www.recordedfuture.com/active-cyber-defense-part-2/ Accessed: 27.04.2018.

[48] Robert M. Lee. Leveraging Cyber Threat Intelligence in an Active Cyber Defense. SANS DFIR Summit. Available at: https://www.youtube.com/watch?v=ea50SyPBDBo Accessed: 13.04.2018, 2016.

[49] Robert M. Lee. Knowing When to Consume Intelligence and When to Generate It. SANS Digital Forensics and Incident Response: CTI Summit, 2017.

[50] Robert M. Lee, 2018. Tweet. Available at: https://twitter.com/RobertMLee/status/996559611133071361 Accessed: 13.05.2018.

[51] Robert M. Lee and David Bianco. Generating hypotheses for successful threat hunting. White paper, SANS Institute Information Security Reading Room, August 2016. Available at: https://www.sans.org/reading-room/whitepapers/threats/generating-hypotheses-successful-threat-hunting-37172 Accessed: 26.03.2018.

[52] Robert M. Lee and Rob Lee. The Who, What, Where, When, Why and How of Effective Threat Hunting. White paper, SANS Institute Information Security Reading Room, February 2016. Available at: https://www.sans.org/reading-room/whitepapers/analyst/who-what-where-when-effective-threat-hunting-36785 Accessed: 26.03.2018.

[53] Lawrence Leung. Validity, reliability, and generalizability in qualitative research. *Journal of family medicine and primary care*, 4(3):324, 2015.

[54] Anna-Karin Lindberg, Sven Ove Hansson, and Carl Rollenhagen. Learning from accidents–what more do we need to know? *Safety Science*, 48(6):714–721, 2010.

[55] Maria B Line, Eirik Albrechtsen, Martin Gilje Jaatun, Inger Anne Tøndel, Stig Ole Johnsen, Odd Helge Longva, and Irene Wærø. A structured approach to incident response management in the oil and gas industry. In *International Workshop on Critical Information Infrastructures Security*, pages 235–246. Springer, 2008. Available at: https://link.springer.com/chapter/10.1007/978-3-642-03552-4_21 Accessed: 22.04.2018.

[56] Maria Bartnes Line and Eirik Albrechtsen. Examining the suitability of industrial safety management approaches for information security incident management. *Information & Computer Security*, 24(1):20–37, 2016.

[57] Lockheed Martin. *The Cyber Kill Chain.* Available at: https://www.lockheedmartin.com/us/what-we-do/aerospace-defense/cyber/cyber-kill-chain.html Accessed: 26.03.2018.

[58] Miroslav Maj, Roeland Reijers, and Don Stikvoort. Good practice guide for incident management, 2010.

[59] Yogesh Malhotra. Organizational Learning and Learning Organizations: An Overview. Available at: http://www.brint.com/papers/orglrng.htm Accessed: 06.05.2018, 1996.

[60] Anders Nese. Improving security posture by learning from incidents. Department of Information Security and Communication Technology, 2017.

[61] Briony J Oates. *Researching information systems and computing*. Sage, 2005.

[62] Chris Pace. Turbocharge Your Threat Hunting Capability With Intelligent TTP Alerting, 2016. Available at: https://www.recordedfuture.com/intelligent-ttp-alerting/ Accessed: 27.04.2018.

[63] Sonal Patel. Threat Actor Behind Cybersecurity Attacks Targeting Safety Instrumented Systems Identified, May 2018. Available at: http://www.powermag.com/threat-actor-behind-cybersecurity-attacks-targeting-safety-instrumented-systems-identified/ Accessed: 25.05.2018.

[64] Michael Quinn Patton. Enhancing the quality and credibility of qualitative analysis. *Health services research*, 34(5 Pt 2):1189, 1999.

[65] Mohammad Farhad Peerally, Susan Carr, Justin Waring, and Mary Dixon-Woods. The problem with root cause analysis. *BMJ Qual Saf*, 26(5):417–422, 2017. Available at: http://qualitysafety.bmj.com/content/26/5/417?etoc= Accessed: 13.04.2018.

[66] The Associated Press. Company: Industrial Hacking Group Has Targets Beyond Mideast, May 2018. Available at: https://apnews.com/75caeec08e1c4dcbaa5e7269efbb3384 Accessed: 29.05.2018.

[67] Scott J. Roberts. Incident Response is Dead...Long Live Incident Response, 2015. Available at: https://medium.com/@sroberts/incident-response-is-dead-long-live-incident-response-5ba1de664b95 Accessed: 27.05.2018.

[68] Scott J. Roberts. Intelligence Concepts – F3EAD, 2015. Available at: https://medium.com/@sroberts/intelligence-concepts-f3ead-964a0653be13 Accessed: 28.04.2018.

[69] Scott J. Roberts. Intelligence Concepts – The Intelligence Cycle, 2015. Available at: https://medium.com/@sroberts/intelligence-concepts-the-intelligence-cycle-f25ec067f1d6 Accessed: 27.04.2018.

[70] Scott J. Roberts. Homemade Ramen & Threat Intelligence: A Recipe for Both. SANS CTI Summit. Available at: https://youtu.be/0lmOKHDBgtc Accessed: 13.05.2018, 2018.

[71] Scott J Roberts and Rebekah Brown. *Intelligence-Driven Incident Response: Outwitting the Adversary*. O'Reilly Media, Inc., 2017.

[72] George Santayana. *The Life of Reason: Vol 1: Reason in Common Sense*. Dover Publications Incorporated, 1905.

[73] Piya Shedden, Atif Ahmad, and AB Ruighaver. Organisational learning and incident response: promoting effective learning through the incident response process. In *Proceedings of the 8th Australian Information Security Mangement Conference*. School of Computer and Information Science, Edith Cowan University, Perth, Western Australia, 2010. Available at: http://ro.ecu.edu.au/cgi/viewcontent.cgi?article=1098&context=ism Accessed: 13.04.2018.

[74] Joe Slowik. Indicators and Network Defense, May 2018. Available at: https://pylos.co/2018/05/16/indicators-and-network-defense/ Accessed: 20.05.2018.

[75] Tove Thagaard. Systematikk og innlevelse. *En innføring i kvalitativ metode*, 4, 2013.

[76] David R Thomas. A general inductive approach for analyzing qualitative evaluation data. *American journal of evaluation*, 27(2):237–246, 2006. Available at: http://journals.sagepub.com/doi/pdf/10.1177/1098214005283748 Accessed: 06.05.2018.

[77] Aksel Tjora. *Kvalitative forskningsmetoder i praksis*. Gyldendal akademisk, 2017.

[78] U.S. Department of Defense. *Joint Publication 2-0: Joint Intelligence*, 2013. Available at: http://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp2_0.pdf Accessed: 27.03.2018.

[79] Johan Van Niekerk and Rossouw von Solms. Organisational learning models for information security. In *The ISSA 2004 Enabling Tomorrow Conference*, volume 30, 2004.

[80] Chris Velazquez. Detecting and Preventing Attacks Earlier in the Kill Chain . Technical report, SANS Institute Information Security Reading Room, August 2015. Available at: https://www.sans.org/reading-room/whitepapers/infosec/detecting-preventing-attacks-earlier-kill-chain-36230 Accessed: 22.04.2018.

[81] Robert K Yin. *Case study research*. Sage, 2009.

Appendix A
Literature Research Method

A literature study was conducted at the beginning of the project to get a thorough understanding of the area of research. Relevant literature from academic institutions and industry organisations was reviewed. Although many industry organisations provide contributing white papers and research, it can be challenging to find these contributions among vast amount of marketing materials and white papers selling the organisation's tools and services. Google Scholar[1], Oria[2] and ScienceDirect[3] were the primary search engines for finding relevant literature. This search engine covers scholarly literature across multiple disciplines and publishing formats. To find related work, a breadth-first search approach was taken, where a set of keywords[4] were used to search for papers in Google Scholar. For each keyword, additional search words were added to scope the search results. Keywords were also combined to find papers related to both. The papers were selected based on the author(s), number of citations, research questions, name of journal or publisher, publication date and research methodology. This selection process included too many papers, and hence a filtering process was needed to reduce the number of papers reviewed. The filtering process used is described below. The papers found in the breadth-first search were read once to extract and document key information. This the following information was noted for each paper:

- Author(s)
- Name of journal and/or publisher
- Year of publication
- Number of Citations

[1] https://scholar.google.com
[2] http://oria.no/
[3] https://www.sciencedirect.com/
[4] Incident Response, Incident Management, Threat Hunting, Threat Intelligence, Cyber Threat Intelligence, Learning from Incidents (LFI)

- Key findings and discussions
- Main conclusions

This structure made it easier to determine which papers should be included in the literature review. The objectives of the final selection process were to assure diversification of the collected papers and avoid a biased scope of the literature. This process was based on human judgement and revised based on feedback on a preliminary project [60].

In addition to reviewing academic and industry literature, standards and guidelines on incident response were studied to gain knowledge on best practices within incident management. These practices influence the data and lessons learned recorded about past incidents. The preliminary project [60] discussed five relevant standards and guidelines:

- NIST Special Publication 800-61, revision 2: Computer security incident handling guide [16]
- ISO/IEC 27035 Standard - Information security incident management [34]
- SANS: Incident Handler's Handbook [40]
- ISF - You Could Be Next [33]
- ENISA - Good Practice Guide for Incident Management [58]

The guide from National Institute of Standards and Technology (NIST) and the standard from ISO/IEC were assessed as the most relevant for this project. The guide from NIST is adopted in wide range of organisations, and the life-cycle introduced in the guide was also found to be a good starting point for a work-flow model integrating threat hunting and information sharing with incident response operations. Of the four guides and standards listed above, all guidelines expect the standard from ISO/IEC are developed by single organisations. Additionally, these practices act as an interface for extensions to incident response procedures in organisations. The guide on computer security incident handling from NIST [16] and relevant standards from ISO/IEC [34], and their relevance for this thesis, are described in Appx. B.

Appendix B

Incident Response Standards and Guidelines

Multiple standards and guidelines describe and recommends procedures for incident response. Two standards (ISO/IEC 27035 [34] and National Institute of Standards and Technology (NIST) Special Publication 800-61 [16]) and three guidelines (SANS: Incident Handler's Handbook [40], ENISA Incident Management [58] and ISF - Learning from incidents to improve resilience [33]) are explored in this section.

B.1 NIST Computer Security Incident Handling Guide

NIST Computer Security Incident Handling Guide provides a guide on computer security incident handling, and is part of NIST Special Publications series on Computer security. Content in this subsection is, if not specified otherwise, derived from [16]. The publication's purpose is to aid organisations in mitigating risks from computer security incidents by providing guidelines on how to respond to incidents efficiently and effectively. The guide is organised in three parts; organising incident response capabilities, phases of handling an incident, coordination, and information sharing.

The first part describes, on a high-level, how an organisation should prepare for an incident. This includes setting up an Incident Response Team (IRT), defining policies to guide incident responders, identifying stakeholders, and ensure that a communication plan is well-defined and ready to be used in the case of an incident.

The second part describes the life cycle of incident response. The four phases of incident response defined by NIST are; preparation, detection and analysis, containment, eradication and recovery, and post-incident activity. The relationship between these phases are illustrated in Fig. B.1.

Preparation This phase involves setting up an IRT, training its team members, deploying systems to help detect and respond to incidents, and identify what is *normal*[1] in an organisation. Implementing mechanisms for risk mitigation is also

[1] An analyst or an Intrusion Detection System needs to know what is normal to be able to detect

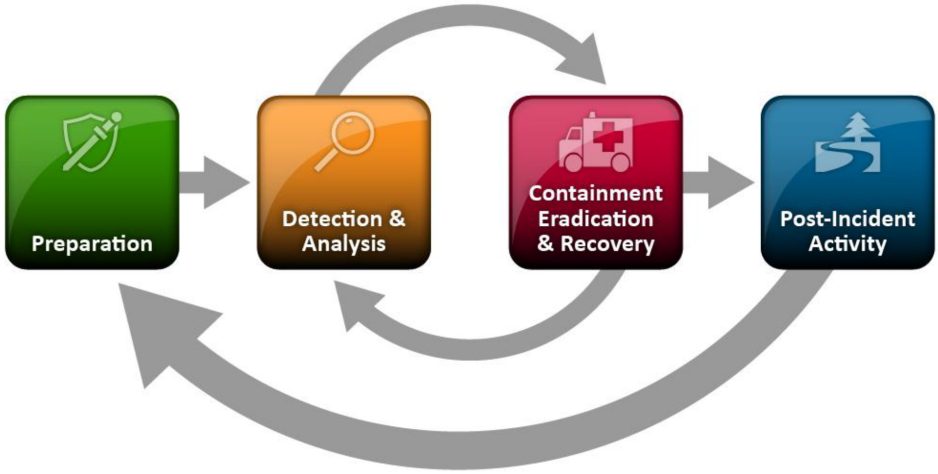

Figure B.1: Incident Response Life Cycle, taken from Cichonski et al. (2012) [16]

part of the preparation phase, but the IRT is not necessarily responsible for this part of the preparation.

Detection and Analysis Organisations should be prepared to detect and handle any incident, and the guideline defines two types of signs that can be used to detect an incident; precursors[2] and indicators[3]. The presence of a precursor or an indicator will require an analysis to determine if this is in fact an incident, or if it is a false positive. Should it be determined that it is an incident, a policy for categorising the incident should aid the analyst give it a proper classification.

Containment, Eradication and Recovery Predetermined strategies and procedures for containing an incident will ease decision-making (isolate host, shut down server, disable user account, etc.) related to an incident. Containment strategies will vary depending on the type of incident, and hence a range of strategies should be predetermined. Eradication could be necessary in some incidents, like malware-based incidents or breached user accounts. It is important to identify all affected hosts and user accounts in such incidents, and ensure that the eradication procedure has successfully completed before starting the recovery procedure. Recovery consists of getting system back online, restoring services and enable disabled user accounts. Both eradication and recovery are OS specific and hence the guideline does not

abnormal behaviour in networks and on hosts. This could be baselining network traffic, common application usage, login activities, etc.

[2] *«A precursor is a sign that an incident may occur in the future»* [16]

[3] *«An indicator is a sign that an incident may have occurred or may be occurring now»* [16]

include any concrete recommendations.

Post-Incident Activity After an incident has been dealt with, the organisation should perform some post-incident activities. Creating follow-up reports and holding lessons learned meetings after major incidents are the only points included in the publication's Incident Handling Checklist, but other activities are encouraged as well. The publication recommends three activities; lessons learned, using collected incident data, and evidence retention. The outcome of these activities should be fed back into the preparation planning to reduce the risk of similar incidents in the future. For instance, if a lessons learned meeting reveals that the root cause of the incident was an unpatched server in the DMZ, then this vulnerability could be mitigated in the preparation phase to avoid it being exploited in the future. Lessons learned should include, but not limited to, determine the root cause of an incident, identify missing information during the response which could have improved the handling of the incident, evaluate if communication and information shared with external parties could have been done better. Use of collected incident data could be to aggregate metrics about single incidents which could be used to evaluate an IRT's operations, improve risk assessments and increase security awareness throughout the organisation.

The third part the publication advises how organisations could ensure effective coordination and information sharing between, and within, their incident response teams and appropriate partners. The publication stresses that it important to clearly define what type of information should be communicated with partners. Information sharing could be done in an Ad Hoc manner, where email, instant messaging and phones are used to share information with peers and coordinate strategies for incident response in a cost-effective way. Cross-organisational coordination and information sharing could also be made partially automated by exchanging information in machine readable format. Such an exchange requires that both parties agree upon a common format on the information. The publication does not recommend a specific format to use, but highlights the need to use a secure transport protocol (like HTTPS, SSL, etc.).

B.2 The ISO/IEC 27035 Standard

The ISO/IEC 27035 Standard is part of the ISO/IEC 27000 family of information security standards [34]. It is currently organised in two parts: *Principles of incident management* (Part 1) and *Guidelines to plan and prepare for incident response* (Part 2). The first part gives a basic overview of incident response, defines definitions and terms, and introduces a phase-based approach of how to handle information security incidents. The second part of the standard gives a more detailed overview of how an organisation can prepare and plan for information security incidents, and how

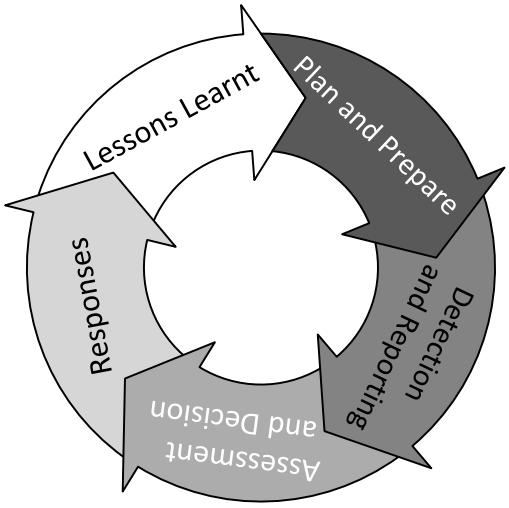

Figure B.2: The five phases defined in ISO27035 [34]

incidents can be used to improve the security posture by learning form the incidents through *Lessons Learnt* activities. ISO/IEC has a similar set of phases as NIST, but splits the two middle phases in NIST's life cycle into three phases as seen in Fig. B.2.

B.3 SANS: Incident Handler's Handbook

The Incident Handler's Handbook from SANS provides information for IT-professionals and managers on how to build an incident response capability by creating incident response policies, standards and teams [40]. This is done in six sequential phases; preparation, identification, containment, eradication, recovery, and lessons learned. The phases in this standard has almost a one-to-one mapping with the phases in the previous two standards, and the lessons learned phases, as in the two others, highlights the importance of using incidents as aids to improve the performance of the IRT and to lay the grounds for improved incident handling in the future.

B.4 ENISA - Good Practice Guide for Incident Management

The Practice Guide for Incident Management by ENISA aims to assist organisations [58]. The guide does not define phases of incident response life cycle, but rather refers to ITIL (IT Infrastructure Library) incident life cycle, which consists of the

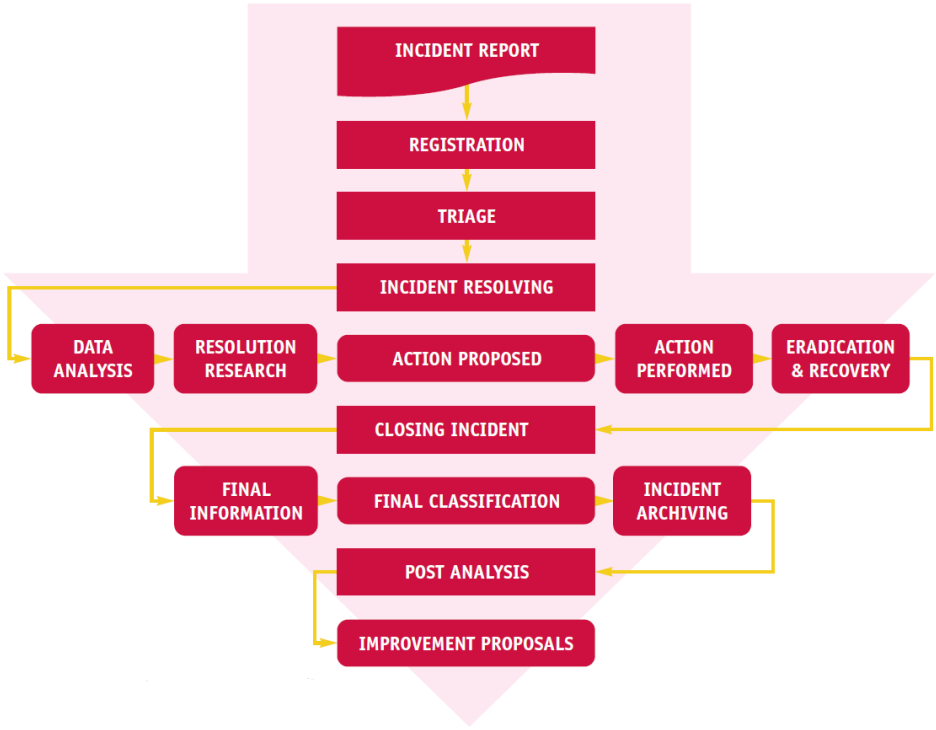

Figure B.3: Incident handling workflow, taken from ENISA's Good Practice Guide for Incident Management [58]

following phases; occurrence, detection, diagnostics, repair, recovery, restoration, and closure [58]. It is worth noting that although this life cycle does not include a separate phase for Learning from Incidents (LFI), it is clearly included in the guideline's Incident handling workflow. For this project, the most interesting steps in this workflow is; Incident Archiving, Post Analysis, and Improvement Proposals.

B.5 ISF - You Could Be Next

'You Could Be Next' is a report from the Information Security Forum (ISF) intended for Chief information security Officers (CISO), security managers and risk managers. It gives recommendations for how to ensure that organisations' incident management process is continuously improving information security by decreasing the likelihood of future incidents, reducing the impact of incidents, and increasing the overall resilience [33]. The ISF report concludes that post-incident reviews empower organisations to improve their response time and develop resilience to withstand

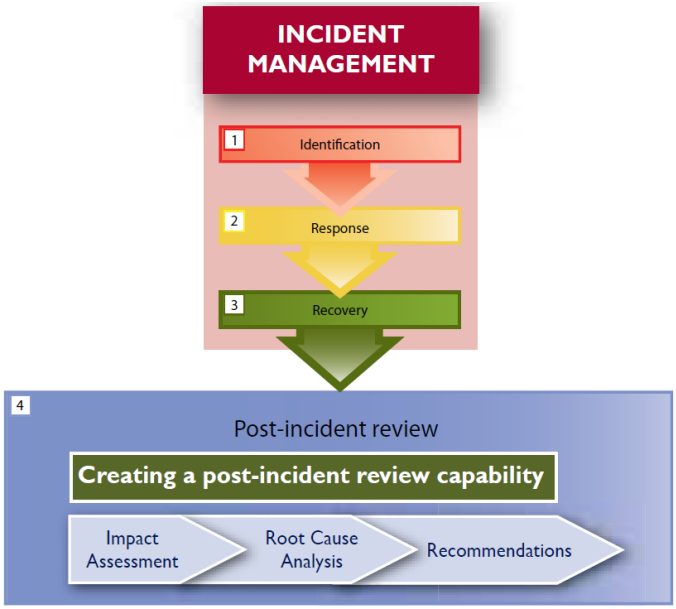

Figure B.4: The ISF Information Security Incident Management process, taken from [33]

impacts from complex threats.

To achieve this, a four-step process to follow when managing an incident is recommended (See Fig. B.4). Note that each step in the workflow expects the previous step to be completed. Creating a post-incident review capability is done by defining a set of impact types and measures, and creating a policy to help determine when a post-incident review should be conducted. The next step in this Post-incident review process is to do an Impact Assessment, which is done by collecting, analysing and communicating the impacts of the incident.

The guide differentiates between a detailed and non-detailed impact assessment, and suggests a method for determine which one to use for a given incident. The next step, finding a root cause of the incident, is optional. Again, a method for determining if the step should be done is suggested by the report. Should the method determine that a root cause analysis should be done, a team to do the job needs to be identified, the analysis conducted, and the outcome of the analysis needs to be communicated to relevant stakeholders. The last step is recommendations, where recommendations are identified and fed in to internal Threat Management and Risk Assessment processes. A new set of incident simulation scenarios should also be created to reveal any gaps in policy or awareness [33].

Appendix C
Interview Guide

In this thesis, we will look at a subtype of incidents called intrusions, and how organisations can improve their security posture by proactively handle and investigate all intrusions, including failed intrusion attempts. We denote intrusions, as in the technical report of the Diamond Model, to "all malicious and nefarious activity targeting computer systems and networks".

The research questions of the thesis are as follows:

- How can intrusion analysis help expedite prevention and detection of intrusions?
- How can indicators be used to discover previously undetected intrusions?

In order to answer the research questions, information about how organisations leverage lessons learned after intrusions will be gathered. Experiences from a variety of intrusions will be systematised through a case study, semi-structured interviews with industry experts will conducted, and a study of how organisations can better utilise lessons learned after intrusions will be performed.

Introduction

1. What is your organisation's core business areas?
2. What is your role within the business?
3. Are you involved in Incident Response in-house, with clients, or both?
4. How is incident response structured in your organisation?
5. Is your organisation using a standard or guideline on incident response, like NIST, SANS, ISO/IEC or ENISA?

General

6. How do you define an incident? How do you define an intrusion?
7. How do you classify intrusions? (severity, successful/failed, targeted, ...)
8. How do you link/correlate related intrusion?
9. Do you see most "targeted" intrusions being part of a campaign, or as stand-alone operations?
10. How are intrusions most commonly detected?
11. When and how do you mark an intrusion as handled?

Documentation

12. How do you document successful intrusions?
13. How do you document failed intrusions?

Lessons Learned

14. Is your organisation doing lessons learned after intrusions?
    a) When is this done?
    b) How is this done?
    c) Who participates?
15. What does your organisation do with the knowledge generated from lessons learned sessions?
    a) How do you assess whether this knowledge is still valid? Tactics, Techniques and Procedures (TTPs) for instance.
16. Is your organisation analysing intrusions (successful and/or failed) to improve its prevention, detection and/or discovery capabilities?
17. How do you think organisations should use lessons learned from intrusions to improve its security posture?
18. How can organisations best structure their utilisation of lessons learned from intrusions?

a) Which challenges would you expect to encounter implementing best practices?

Retrospective Analysis

19. Is your organisation analysing past intrusions in order to discover undetected related intrusions, for instance failed intrusion attempts prior to the successful one?

    a) Are you applying any particular method? E.g. Root cause analysis, structured intrusion analysis, etc.

20. What knowledge from successful intrusions is of particular interest during the retrospective analysis? (IP addresses, domains, file names, email subjects, TTPs, targets, . . .)

Threat Hunting

21. How do you incorporate knowledge about your organisation when generating threat hunting hypotheses?

22. How do you incorporate knowledge about your organisation during threat hunting operations?

23. What potential do you see in using lessons learned as input to generating threat hunting hypotheses?

24. What are the main challenges in implementing efficient threat hunting operations?

25. How can intrusion analysis enable and improve proactive detection of unknown intrusions?

Sharing Information

26. Is your organisation sharing lessons learned with external parties?

    a) What type of information do you mainly share (IP addresses, domains, file names, email subjects, TTPs, targets, . . .)

27. What type of information/indicators? have you received that has been of particular interest/help? (IP addresses, domains, file names, email subjects, TTPs, targets, . . .)

28. What potential do you see in sharing lessons learned with external parties, like cooperating organisations, partnerships, national agencies, etc.?

    a) From the perspective of the sharing organisation

    b) From the perspective of the receiving organisation

29. Is cooperation with external parties (agencies, vendors, partners, Computer Emergency Response Teams (CERTs), etc.) and/or Threat Intelligence helpful for detecting and/or preventing intrusions?