# THE STUDY OF THE OPERATIONAL PRACTICES OF NATIONAL CSIRTS REGARDING THE USE OF FREE TOOLS AND PUBLIC DATA IN SUPPORTING COMPUTER SECURITY INCIDENT RESPONSE

A THESIS SUBMITTED TO
THE UNIVERSITY OF KENT
IN THE SUBJECT OF COMPUTER SCIENCE
FOR THE DEGREE
OF PHD.

By
Sharifah Roziah Binti Mohd Kassim
02 July 2023

# Abstract

Computer Security Incident Response Teams (CSIRTs) have been established at national and organisational levels to respond and coordinate responses to computer security incidents. It is known that many CSIRTs, including national CSIRTs, routinely use different types of tools and data, including free tools, open-source tools, and public data, in their daily work to support incident responses. However, a lack of public information and systematic discussions was observed regarding how national CSIRTs use and perceive free tools and public data in their operational practices.

To achieve a more comprehensive and systematic understanding of how national CSIRTs use and perceive free tools and public data in the operational practices, a systematic literature review (SLR) of the research literature and websites of national CSIRTs and cross-CSIRT organisations was conducted in this research. A primary finding from the SLR is that most discussions concerning free tools and public data used in national CSIRTs' operations are largely incomplete, ad hoc, and fragmented. This includes a lack of discussions on how the staff of national CSIRTs perceive the usefulness of free tools and public data to facilitate incident responses. Such gaps can prevent us from understanding how national CSIRTs can benefit from free tools and public data and how other organisations, individuals and researchers can help by providing such tools and data to improve national CSIRTs' operations. The findings from the SLR call for more empirical research on how national CSIRTs use and perceive free tools and public data. This should also include how such data and tools can be leveraged to support incident responses at national CSIRTs' operations.

Hence, a survey and twelve follow-up semi-structured interviews with staff members of thirteen national CSIRTs worldwide were undertaken in this research to gain insights into how free tools and public data are used and perceived in national CSIRTs. The study was conducted in two phases: first, with staff members of the Malaysia Computer Emergency Response Team (MyCERT) to get some initial results, and then with twelve other national CSIRTs to enlarge the results from the first phase. Results from the survey and the semi-structured interviews led to three main findings: 1) confirmation from the participants regarding the active use of free tools, public data, and open-source intelligence (OSINT) tools in national CSIRTs, 2) the perceived usefulness of free tools and public data to support incident responses in national CSIRTs, and 3) a lack of systematic procedures in guiding the use of free tools and public data across the participating national CSIRTs (for example, one aspect is on how such tools and data should be evaluated for quality and usability). The finding on the lack of systematic procedures for evaluating free tools and public data calls for further research and development to better understand current tools and data evaluation practices in national CSIRTs. Such understanding shall inform researchers on developing systematic procedures for evaluating free tools and public data in national CSIRTs' operations.

An empirical study using several focus group discussions was conducted to understand better current tools and data evaluation practices in real-world operations of national CSIRTs. The findings from the focus group study confirmed that the evaluation practices are ad hoc and informal. Systematic procedures that leverage industry standards, such as criteria for evaluating free tools and public data, are unavailable in the participating national CSIRTs' operations. This finding informed the construction of candidate criteria for evaluating free tools and public data from the focus group discussions. Nevertheless, the validity of the candidate criteria for usefulness, deployment and applicability is uncertain. This calls for studies to validate the candidate criteria before implementation in real-world operational practices of national CSIRTs for evaluating free tools and public data.

A validation study using semi-structured interviews was conducted to gain insights into how staff members of national CSIRTs perceive the usefulness and deployment of the candidate criteria in national CSIRTs. This is followed by a more objective validation by applying the criteria to evaluate two candidate tools and one sample data source for applicability in practice. This was performed by converting each criterion into one or more relevant metrics, such as "measuring the time taken by a tool to produce results". Significantly, results from both validation approaches were consistent, leading to the following findings: 1) the candidate criteria were perceived as practically useful for evaluating free tools and public data in the operations of national CSIRTs; 2) the candidate criteria were perceived as ready for deployment in national CSIRTs and 3) the criteria is applicable in practice to evaluate free tools and public data. It is envisaged that these criteria would help national CSIRTs and the broader security operations to select usable and good quality free tools and public data available on the Internet to support incident response. Subsequently, enhancing the current practices in evaluating free tools and public data in national CSIRTs.

# Acknowledgements

First and foremost, I thank God, the Almighty, for giving me boundless care, excellent health, motivation, patience and a conducive environment to complete this thesis as scheduled.

I thank my husband (Othman) and my children (Muhammad Marwan, Muhammad Naim and Muhammad Ahnaf) for their continuous understanding, support and patience throughout my PhD journey. I also would like to thank my mother and siblings for their support and encouragement.

I sincerely thank Professor Shujun Li and Dr Budi Arief for their excellent guidance, expert knowledge, supervision, motivation, encouragement, and great insight throughout this PhD's journey.

I thank the Malaysia Computer Emergency Response Team (MyCERT) staff members for their support in participating in my research projects during my PhD. I want to thank the CEO of CyberSecurity Malaysia, Dato' Ts. Dr Haji Amirudin Bin Abdul Wahab and the CTO, Dr Solahuddin Bin Shamsuddin, for their support and motivation to take up the PhD.

Last, I am thankful to all my colleagues, friends and staff members of the University of Kent, United Kingdom, who helped me when I needed assistance and support during my PhD.

# List of Relevant Outputs from the PhD Research

## Published Peer-reviewed Research Papers

1. Sharifah Roziah Binti Mohd Kassim, Shujun Li and Budi Arief (2022). How National CSIRTs Leverage Public Data, OSINT and Free Tools in Operational Practices: An Empirical Study. *Cyber Security: A Peer-Reviewed Journal*, 5(3):251–276. `https://www.ingentaconnect.com/contentone/hsp/jcs/2022/00000005/00000003/art00007`

2. Sharifah Roziah Binti Mohd Kassim, Shujun Li and Budi Arief (2022). Incident Response Practices Across National CSIRTs: Results from an Online Survey. *OIC-CERT Journal of Cyber Security*, 4(1):67–84. `https://www.oic-cert.org/en/journal/vol-4-issue-1/5.html`

3. Sharifah Roziah Binti Mohd Kassim, Solahuddin Bin Shamsuddin, Shujun Li and Budi Arief (2022). How National CSIRTs Operate: Personal Observations and Opinions from MyCERT. In: *Proceedings of the 2022 IEEE Conference on Dependable and Secure Computing* (IEEE DSC 2022). 4(3):1–2. `https://doi.org/10.1109/DSC54232.2022.9888803`. (Won the *Best Paper Award* for the Experience and Practice Track)

4. Sharifah Roziah Binti Mohd Kassim, Shujun Li and Budi Arief (2023). Understanding How National CSIRTs Evaluate Cyber Incident Response Tools and Data: Findings from Focus Group Discussions. In *ACM Digital Threats: Research and Practice (DTRAP)*, 4(3):45:1–45:24. `https://doi.org/10.1145/3609230`

## Conference Presentations

1. 19th October 2022: Incident Response Practices Across National CSIRTs: Results from an Online Survey. Presented virtually as an oral talk at the 20th Annual APCERT Annual Conference.

2. 7th September 2022: Cyber Incident Response Practices Across National CSIRTs' Operations: Results from an Online Survey. Presented as a poster at the 24th International Conference on Information and Communications Security (ICICS 2022), Canterbury, UK.

3. 1st July 2022: Use of Public Data, OSINT and Free Tools in National CSIRTs: Findings from an Empirical Study. Presented in person as an oral talk at the 17th National CSIRTs Meeting, Dublin, Ireland.

4. 30th June 2022: Use of Public Data, OSINT and Free Tools in National CSIRTs: Findings from a Systematic Literature Review and Empirical Study. Presented in person as an oral talk at the 37th Annual FIRST Conference in Dublin, Ireland.

5. 24th June 2022: How National CSIRTs Operate: Personal Observations and Opinions from MyCERT. Presented virtually as an oral talk at the 2022 5th IEEE Conference on Dependable and Secure Computing (IEEE DSC 2022) in Edinburgh, UK.

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

Chapter 1 of the thesis introduces the research and its relevance to the cyber security field. It presents the state of the cyber security threat landscape, providing context for undertaking the research. This chapter is divided into six sections. Section 1.1 briefly describes the current cyber security threat landscape, its impacts on nations, organisations and citizens, and how this research fits in. Section 1.2 explains the reasons behind undertaking the research. Section 1.3 describes the aims and research questions (RQs), while Section 1.4 briefly describes the methodology adopted for the research, including data collection, data analysis and research ethics review. Section 1.5 provides the contributions of the research. Finally, Section 1.6 outlines the organisation of this thesis, consisting of eight chapters, including Chapter 1.

## 1.1 Cyber Security Current Threat Landscape

Due to incredible convenience and fast communication, the Internet has become essential for governments, industries, businesses and citizens to conduct various activities online, from business transactions to social networking. At the same time, cyber attacks are also taking place by exploiting vulnerable systems across networks, and human weaknesses, taking advantage of borderless cyber space without jurisdiction for malicious purposes [246].

Worldwide cyber attacks are increasing in scale and rapidity, and their impact and losses are becoming challenging to predict [96]. This can be seen from cyber attacks expanding globally in 2021, compared to 2020 [264]. In 2021, the average number of cyber attacks and data breaches increased by 15.1% from the previous year [34]. The European Union Agency for Cybersecurity (ENISA) Threat Landscape Report 2022 stated that cyber security attacks continued to increase during the second half of 2021 and 2022, not only in terms of vectors and numbers but also in terms of their impact, with ransomware and threats against availability rank at the top during the reporting period [144]. In the

UK, more than 85% of organisations had experienced successful cyber attacks in 2020 and 2021, and over 12 months, ransomware attacks affected 73% of UK organisations [48]. Besides disrupting services and businesses, cyber attacks have caused substantial monetary losses to involved organisations worldwide. The National Fraud Intelligence Bureau (NFIB) Fraud and Cyber Crime Dashboard reported losses to fraud and cybercrime in the UK totalled over £4.1 billion for the past 13 months, in 2022 [192]. According to a 2022 report, a survey among Information Technology (IT) decision-makers in the United States of America (USA) found that nearly a quarter of companies that have experienced a cyber attack in 2022 have lost between 50,000 and 99,999 US dollars [238]. As of July 2022, the average cost of cyber security breaches in the previous 12 months in the UK was £1,100 across all businesses [238]. Pricewaterhouse Coopers, in their 2022 Cyber Threat Report, found companies in 2022 battled with sophisticated advanced persistent threats or APTs, aggressive cyber criminals, disgruntled insiders, hacktivism and distributed denial of service (DDoS) attacks [209].

Today's cyber threat landscape continues to see an increase in cyber criminals of all motivation and skill levels who choose their targets more strategically. This can be seen from threat actors who continuously enhance their tactics and techniques and share their tools, motivated by sabotage, espionage and money [209]. Other motivations of cyber criminals can range from information theft and data destruction to gathering critical information to launch an even more significant impact attack, i.e. a major disruption to the power station, backbone or communications infrastructure, threatening the national security and other national critical services [96]. These attacks can have severe impacts and repercussions on organisations and entities if not responded to and mitigated effectively. For instance, the devastating Wannacry ransomware attack in May 2017 that targeted the UK National Health Service (NHS) caused the cancellation of 19,000 appointments. The cancellation cost the NHS £20 million loss of revenue between May 12, 2020, and May 19 2020 and another £72 million incurred in recovering its IT systems [65]. In another instance, in September 2017, a consumer credit card agency, Equifax, was targeted, where personal data belonging to 147 million Americans (almost 44% of the American population) was breached and exposed [249].

Furthermore, during and since the COVID-19 pandemic [278], people and organisations have become much more reliant on digitally connected devices and services. Such a reliance inadvertently increases people's exposure to cyber threats [14], unexpectedly impacting the cyber security threat landscape during the COVID-19 pandemic [145]. This is due to the expansion of work from home, e-learning and online shopping, leading to unforeseen new vulnerabilities [137; 88], contributing to the increase in cyber attacks worldwide [101]. It is also worth mentioning that the Russian-Ukraine War, beginning in February 2022, has contributed to the rise in cyber attacks, particularly within the European region, to some extent. For instance, state-sponsored hacking groups launched more than 237 cyber attacks against Ukraine in the weeks before Russia's February 24 invasion [254]. This includes 37 destructive malware attacks against Ukraine between February 24 and April 8 aimed at industrial control systems (ICS) in Ukraine. In another instance, a cyber attack disrupted satellite Internet connections across Europe just before the invasion of Ukraine began on February 24 [24]. Apart from the two significant events between 2020 and 2022, delays in applying patches [15] and using outdated systems, [71] generally contribute to the increase in cyber attacks. Lack of qualified tools, data and security experts for responding to cyber attacks compounded the situation by delays and efficiency in responding to cyber attacks [207].

Overall, the current cyber security threat landscape is becoming a serious global issue that must be addressed efficiently. Governments worldwide are concerned that cyber attacks could seriously impact critical national information infrastructure and citizens if they are not responded to and resolved efficiently. To address cyber attacks and their impacts, many countries have made securing critical infrastructures, citizens and public services with preventive and defence measures an integral part of their national cyber security strategies [18; 134]. Notable examples of such national strategies have been defined by governments of the UK [180], the USA [277], Malaysia [153] and Singapore [47].

While preventive measures such as applying security updates, performing backups, and regular network security inspections are essential, it is not sufficient to rely solely on them. Some widely used technologies, such as intrusion detection systems (IDS), cannot respond and handle incidents but only detect and give alerts about possible cyber attacks [92]. Instead, efficient responses to computer security incidents are becoming more critical now [133; 224]. Schneier (2014) has said: *Security is a combination of protection, detection, and response, with the 1990s being the era of protection, while this decade now is the era of response* [224]. In essence, the "response" highlighted by previous researchers in addressing cyber attacks fits into the research reported in this thesis.

It should be noted that the responsibility to handle and mitigate cyberattacks before they become widespread within a nation lies with the government [78]. This responsibility includes the establishment of Computer Security Incident Response Teams (CSIRTs) at the national level to safeguard a nation's cyber security and act as a national cyber security reference and coordination centre [64]. Creating and maintaining CSIRTs, including national CSIRTs at the state or governmental level, to detect, mitigate and recover from computer security incidents is crucial [186]. A study by the CERT Division of the Software Engineering Institute (SEI) of Carnegie Mellon University (CMU) in the USA in 2022 has highlighted five essential components of national CSIRTs that are critical to ensure the sustainability and success of national CSIRTs in responding to and mitigating cyber attacks within a constituency efficiently [27], depicted in Figure 1. A component highlighted by Bills et al. (2022) "establish effective operations by proposing and recommending effective operational practices, i.e.procedure and guidelines to facilitate incident response in national CSIRTs". Taking together the importance of "response" highlighted by Koivunen et al. (2010) and Schneier (2014) with "recommendation of effective operational practices with procedures in national CSIRTs" by Bills et al. (2022) fit into the research reported in this thesis.

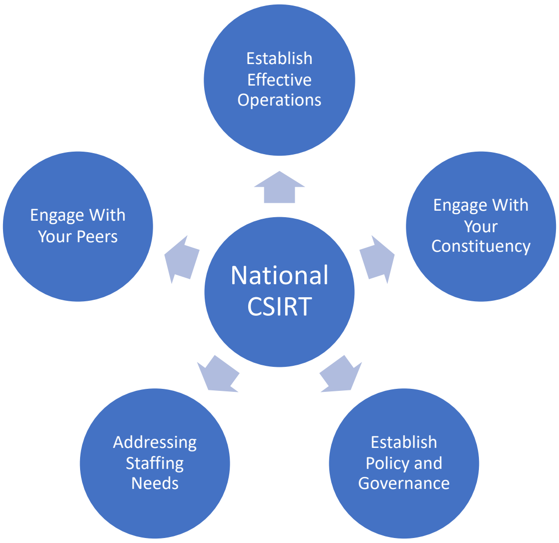

Establish Effective Operations
Engage With Your Peers
Engage With Your Constituency
National CSIRT
Addressing Staffing Needs
Establish Policy and Governance

Figure 1: Components of National CSIRT Establishment [27]

The research reported in this thesis is about better understanding the operational practices of national CSIRT regarding the use of free tools and public data to support incident responses. Such understanding would help to identify any operational gap in the current practice and to construct a procedure that addresses the gap and enhances current practice. Research and development to design cyber security solutions for responding to and mitigating cyber attacks are suggested to minimise cyber attacks' impact. Kijewski and Kozakiewicz (2011) expressed a crucial need for constant research in cyber security to develop measures, including procedures to respond to cyber threats more effectively – supporting the research reported in this thesis [125]. It is timely that this research is undertaken when cyber attacks continue to increase globally with sophisticated attack techniques, demanding a more effective response to cyber attacks, justifying why this research is vital for protecting cyber space.

**Scope of research.** It should be noted that the scope of this research is national CSIRTs. Organisation CSIRTs and other types of CSIRTs are not included in the scope. The fact that national CSIRTs play a key role in safeguarding a nation from cyber attacks and as an authorised national point of contact concerning cyber attacks [91; 173; 289] makes them better study subjects than organisation CSIRTs. The key roles include responding and coordinating responses to cyber attacks [279; 9; 60; 214], and protecting critical national infrastructures from cyber attacks [94; 186; 289]. In addition, national CSIRTs also have received greater attention in cyber security research, critical information infrastructure protection (CIIP), and national crisis management issues in a country [289].

The key role of national CSIRTs has been emphasised from a legal perspective, whereby the Pan-European Cooperation Commission in 2009 requested its member states to ensure that national CSIRTs must be the key component of national capability [63]. This includes preparedness, information sharing, coordination, response and leading national contingency planning and exercises [63]. Furthermore, In 2016, the European Parliament and the European Council passed the EU NIS Directive [64], which requires EU member states to have well-functioning national CSIRTs. This function is to "comply with essential requirements to guarantee effective and compatible capabilities to deal with incidents and risks and ensure efficient cooperation at Union level".

Another justification is as the researcher was collecting empirical evidence from the real world for this research, the scope of the research was limited to national CSIRTs based on "access" as the researcher has built a professional network across national CSIRTs. Notably, gaining access to sensitive information related to cyber operations is very difficult to secure at organisations CSIRTs [275]. Nevertheless, the researcher has networks with national CSIRTs that can make such rare access to information at national CSIRTs possible. In contrast, the researcher does not have such access to a good sample of organisational CSIRTs.

**Terminology.** It should be noted the focus of this research is on "free tools" (including "open-source tools", "free online services" and "free OSINT tools") and "public data". Throughout this research, the following terms are used to define the scope and focus of the research more precisely:

1. Free tools: software tools and online services available for free to help facilitate cyber incident responses. This includes open-source tools.

2. Public data: data available to the general public for free without restrictions, such as regarding sharing and disseminating the data.

3. Open-source Intelligence (OSINT) tools: free software tools and online services explicitly used to facilitate and extract cyber threat intelligence (CTI) for incident responses.

4. Open-source tools: a significant class of free tools, referring to free tools whose developers publish the source code for anyone to access.

5. Free online service: refers to free web-based software tools.

6. National CSIRT staff: employees of national CSIRTs who are involved in the daily handling of incidents in the operations, such as analysts, team leaders and executives who know how incidents are responded to in their operations.

7. Cyber incidents: any cyber event having an actual adverse effect on the security of the network and information systems.

8. Cross-CSIRT organisations: organisations that help connect different CSIRTs. The cross-CSIRT organisations referred to in this thesis are the Forum of Incident Response and Security Team (FIRST)(`https://www.first.org`), the International Telecommunication Union (ITU)(`https://www.itu.int`), the Software Engineering Institute (SEI) of the Carnegie Melon University of USA(`https://www.sei.cmu.edu`) and the European Network Information Security Agency (ENISA)(`https://www.enisa.europa.eu`).

Background information on incident response, CSIRTs, national CSIRTs and other related terms that need to be known to understand better the research reported in the thesis is provided in Chapter 2.

## 1.2 Motivation

My motivation to undertake this research came from my real-world experience and personal observations while working at a national CSIRT, the Malaysian Computer Emergency Response Centre (MyCERT) [167].

From my observation, staff members at MyCERT rely on data from various sources and a wide range of tools to facilitate incident responses, particularly free tools, open-source tools and public data. I observed that free tools, open-source tools, free online services, and public data are ad hoc. This gave me the impression that standardised procedures for guiding better use of such tools and data were lacking. Moreover, systematic and open discussions on how free tools, open-source tools and public data are used and perceived at national CSIRTs were also lacking. Some free tools and public data sources may not have been utilised sufficiently, potentially due to a lack of systematic information about such tools and data and how to search them easily.

While working at MyCERT, I also observed that procedures to evaluate and choose which tools and data to use for incident response were lacking. This is particularly a problem for free tools and public data, which often do not go through a proper quality assurance process like commercial tools and data. In contrast, I observed that a formal evaluation was usually practised for commercial tools and data as part of standard procurement procedures. I believe methods for systematically evaluating tools and data are essential for national CSIRTs to specify the quality of tools and data for incident response, echoing results of previous study [32].

In a nutshell, my experiences and personal observations while attached to MyCERT motivated and inspired the research reported in this thesis in two key areas [167]:

1. To better understand the operational practices of national CSIRTs concerning the use of free tools, open-source tools and public data and how these can be used better by national CSIRTs.
2. To enhance current practices at national CSIRTs by developing systematic procedures for key areas of operational practices – one such area is tool and data evaluation.

It is a noble intention to understand better the operational practices across national CSIRTs and enhance the current practices, precisely in critical areas of operations. The research reported in this thesis represents a vital component of the overall cyber security strategy in mitigating cyber attacks and protecting nations' critical information infrastructure and citizens. This is the gist of the research reported in this thesis.

## 1.3 Research Questions

**Research Aim.** This research aims to achieve a better understanding of the operational practices at national CSIRTs. In particular, it will focus on current practices regarding the adoption of free tools and public data to support incident responses with the goal of enhancing the practices.

**Research Question (RQ).** There are four research questions defined as below:

RQ1 What is the state of the art regarding the use of free tools and public data in the operations of national CSIRTs?

RQ2 How are free tools and public data used in the real-world operations of national CSIRTs?

RQ3 What criteria can be used to evaluate the suitability of free tools and public data in incident response practices of national CSIRTs?

RQ4 How can a proposed set of criteria help to evaluate the suitability of free tools and public data in incident response practices of national CSIRTs?

The four RQs are linked to four relevant chapters in the thesis as follows:

- To answer RQ1, a systematic literature review of websites of national CSIRTs, cross-CSIRTs organisations and research papers was undertaken, reported in Chapter 3 under Section 3.5.
- To answer RQ2, a survey and some semi-structured interviews with national CSIRTs were performed to gain insights into how free tools and public data are used in the operations and how staff members perceive the usefulness, reported in Chapter 4.
- To answer RQ3, focus group discussions with national CSIRTs were conducted to identify a set of criteria for evaluating tools and data, reported in Chapter 5.
- To answer RQ4, semi-structured interviews with national CSIRTs were conducted to validate the proposed criteria for usefulness and deployment and apply the criteria to evaluate two tools and a data source widely used by national CSIRTs, reported in Chapter 6.

## 1.4 Research Methodology

The *aim* of this research is to achieve a better understanding of the operational practices at national CSIRTs concerning the use of free tools and public data in supporting incident responses and to come up with a practical procedure to enhance the practices. A qualitative methodology was primarily adopted to achieve the research aim. A minor quantitative methodology was also adopted, mostly around descriptive statistics and quantitative content analysis for counting numbers and translating them into charts, histograms and percentages. Due to the qualitative nature of the overall research and the small number of participants, which is common in empirical research, it did not allow an advanced quantitative methodology. The quantitative analysis in this research provides only an overview and general understanding of the data. Moreover, the research design did not consist of empirical hypotheses tests; hence, an advanced quantitative method was not considered [204]. Notably, qualitative research is a rigorous approach involving extensive time in the field, working in the often complex, time-consuming data analysis process and writing detailed, accurate findings [234].

In principle, qualitative research covers the following research designs: phenomenology – focuses on individual experiences; grounded theory – generates new theory; ethnography – studies the culture of a group of people; and case studies – focus on a specific individual case [162; 205]. Phenomenology is eliminated for this research as Phenomenology studies the lived experiences of individuals in this world and their situation in detail [162]. In contrast, the analysis here is not concerned with exploring the participants' lived experiences but rather inquiring about the participants' knowledge and opinions of the operations of national CSIRTs. The grounded theory approach is inappropriate as this research does not intend to generate a new theory on a particular subject area [162]. A case study is excluded as this research does not involve an intensive study of a single case or provide an in-depth picture of an individual, group or organisation [205]. Ethnography, which focuses on investigating social customs, beliefs and behaviours that define a culture, is eliminated, as the research reported in this thesis is not to study social culture or behaviour of individuals [205]. Looking at all the above research designs, a "generic qualitative inquiry" design was considered for the research, focusing on external and real-world, rather than internal and psychological, as with other qualitative research designs [205; 119]. Generic qualitative studies could incorporate some elements and strengths of established qualitative methodologies while maintaining flexibility. This makes generic approaches more attractive to researchers when their studies do not fit neatly within a particular established methodology [119]. *Inductive* nature of research using semi-structured interviews, focus groups, open codes, categories, thematic and content analysis are most common in the "generic qualitative inquiry" approach [119], which fits into this research.

The research type used for this study is inductive, whereby findings are generated and interpreted from the bottom-up, grounded from the exploratory data, contrary to a deductive research type. A rationale for choosing inductive research over deductive research is the scarcity of literature on this research topic. Therefore, inductive research is necessary to generate as much information from the data as possible to allow new information, themes, concepts, categories, and messages to emerge.

### 1.4.1 Data Collection

Data collection for qualitative inquiry often requires semi-structured interviews, surveys, and focus group discussions on collecting in-depth first-hand data [205]. This section briefly describes the data collection methods used in this research.

#### Survey

A survey was chosen as a data collection method for this research as it is a flexible research tool with a good reputation in the past decade in academia, government, and private sector to gather information from participants coming from various backgrounds for studying multiple fields of study [217]. In this research, a survey was used to collect initial information from participants on the research topic, followed by semi-structured interviews to elicit more detailed information in addition to the survey. A similar approach was adopted in previous studies to gain insightful information about CSIRTs using surveys and followed by semi-structured interviews [82; 97; 116].

Surveys are generally conducted over the phone or through paper-based questionnaires. However, for the convenience of potential participants, this research adopted an online survey, and the instrument used in the study is an online survey questionnaire [217].

The research used a subscription-based online survey platform for security and privacy compliance purposes instead of a free one.

More chapter-specific details of the survey conducted for the research are explained in Chapter 4.

### Semi-structured Interview

A qualitative interview is one of the most essential data-gathering tools in qualitative research [176] and is used to draw insights about a particular topic. Interviews provide the researcher with various options and focus more on the qualitative analysis of the interview data. As this research is exploratory, qualitative interviews (and focus groups) complement this nature by focusing on the "qualities" of opinions from participants to uncover diverse meanings from participants' experiences, aligning with the research aim and objectives [86].

While structured interviews are quantitative and primarily used for hypothesis testing, semi-structured interviews were preferred for this research, which allowed the researcher some flexibility in asking questions and the order of the questions. Unstructured interviews are too flexible, whereby the interviewer provides context for the topic of interest, and the interviewee is just encouraged to talk as they would like about the topic [39]. In this research, an interview method that sits between structured and unstructured was preferred, neither seeking extensive data for testing a hypothesis nor fully letting interviewees provide the context for the topic and drive the interview. Therefore, a semi-structured interview is deemed most appropriate for obtaining qualitative input from interviewees to fulfil the research aim and objectives.

For research purposes, the Internet can potentially be used as a revolutionary approach to collect primary and secondary data through an online survey, online interviews, and online focus groups  [46]. Online interview is also known as "e-interviews", referring to interviews that are conducted using computers and the Internet as the medium or "computer-mediated communication" (CMC), e.g. computers, cell phones or mobile devices, using text chat or messaging, multichannel web conferencing spaces, video conferencing [219]. This research adopted online semi-structured interviews as an alternative option to enable conversation at any time and from anywhere [219], feasible, economical, time-saving with wide geographic coverage [197]. Moreover, with the global COVID-19 pandemic at the time of the research between the years 2020 and 2022 [278], online interviews were seen as the safest means for collecting data through interacting with people. The instrument used for the online semi-structured interviews consists of an "Interview Schedule" that defines the interview guidelines, with interview questions that guide the interview throughout. Interview questions were arranged from general to specific, with some flexibility in the order of how questions are asked during the interviews.

More chapter-specific details of the semi-structured interviews conducted for this research are explained in Chapters 4 and 6.

### Focus Group Discussion

Focus group discussion is a data collection method with several participants from similar backgrounds to obtain rich data through group interaction [141]. It is a well-known data collection method used in many fields of study, including cyber security research for developing procedures, guidelines and other strategies [8; 284].

Focus group discussion was adopted for the research to collect rich information from participants consisting of national CSIRTs' staff members through free-flow group interactions [89]. Focus groups allowed the exploration of "non-sensitive topic" of study through collective open opinions, views, and experiences within a group context, where interviews and surveys are unsuitable for collecting data within a group setting [75; 220]. Focus groups are also suitable for drawing out expert knowledge and experiences from people with expertise in specific domains, which may have a great deal of information, as desired for this research [213]. Hence, focus group discussion was the best fit to obtain rich data from collective expert opinions to achieve this research aim and objectives.

In principle, focus group discussions are conducted face-to-face to gather data for research. This research adopted "online focus group discussion" due to the international nature of the study, where participants are dispersed across regions, coupled with the current global COVID-19 pandemic between the years 2020 and 2022 [278] that limits travelling and face-to-face interaction with people. Therefore, an online focus group was the safest option for collecting data from staff members of national CSIRTs worldwide. Additionally, it saved money and time if travelling was necessary for collecting data in this research. The instrument used for the focus group discussions is a "Focus Group Schedule" containing an agenda and general guidelines with a list of questions arranged sequentially, from general to specific [141].

More chapter-specific details of the focus group discussions conducted for the research are explained in Chapter 5.

## 1.4.2 Data Analysis

In this research, three data analysis methods were used: thematic analysis, content analysis, and descriptive statistics, with thematic and content analysis being the primary methods and descriptive statistics being minor. Each data analysis method used throughout the research is described in this section.

### Thematic Analysis

Thematic analysis, a widely-used data analysis method supporting qualitative research, was used to analyse data collected from some semi-structured interviews reported in Chapter 4. It is primarily built upon the theoretical positions of Braun and Clarke [33]. Thematic analysis is used in this research to explore and identify new themes and patterns that emerge across data, allowing interpretation of its underlying meaning [33]. Thematic analysis was chosen for this research to ensure a rigorous analysis to produce insightful findings that answer the research question [33]. Moreover, thematic analysis is useful when investigating an under-researched research topic demonstrated by the present research [33] and was the best fit to analyse the small data sets collected from 12 interviewees in Chapter 4 of the thesis [266].

The thematic analysis also complimented the exploratory nature of this research through a data-driven inductive coding to generate as many themes from the data [33]. In thematic analysis, a theme captures significant ideas about the data concerning the research aim and represents patterns or meanings within the data set [33]. A theme can be captured and defined from two approaches: “manifest” – which captures the surface meaning, and “latent” – which dives into deeper meaning with underlying patterns [38]. In this research, a combination of both “manifest” and “latent” approaches was adopted to help maximise the research’s scope, analysis, validity and reliability [38].

Nevertheless, the manifest approach was primarily adopted to describe what is directly observed without assigning a deeper interpretation or deeper meaning during analysis [38]. A minor latent approach was adopted to minimally interpret the underlying meaning of the data, focusing on the implied value of the data from the judgement of the researcher [38]. In thematic analysis, the data is analysed by coding [195] to capture significant themes across the data [195]. The codes are descriptive and capture the data’s essence, allowing interpretation of its meaning.

More chapter-specific details of the thematic analysis method used in this research are explained in Chapter 4 of the thesis.

### Descriptive Statistic Analysis

Descriptive statistics were used to summarise the data collected from the survey conducted in this research as presented in Chapter 4. It is an exploratory data analysis method that describes quantitative data from a survey in a simple quantitative form by counting, summarising, organising and presenting it in graphical forms [132]. This helped to provide a high-level understanding of survey data. Descriptive statistics was also preferred for this research to analyse the survey data because it helped collect and summarise data more manageable and organised, making the results easier to comprehend [122]. It is a straightforward process which easily translates survey data into results in percentages and overall averages [122].

A detailed descriptive statistic analysis method used in this research is provided in Chapter 4 of the thesis.

### Content Analysis

Content analysis was used in this research to analyse data collected from the focus groups reported in Chapter 5 and the data collected from semi-structured interviews reported in Chapter 6 to capture participants' opinions and views to answer the corresponding RQs. A content analysis was also used to analyse the research papers, online documents and web pages in the SLR reported in Chapter 3. Content analysis aims to describe, explain data objectively, categorise and quantify specific concepts, words or phrases in the data, for instance, the frequency of a particular word that appears in the data or categorise tools by classes that appear in a data [203]. It focuses on describing and categorising qualitative data objectively rather than identifying themes in the data, as with thematic analysis.

A manifest approach in qualitative data analysis can capture the concepts, significant information and key terms at the surface level of the data without looking for profound interpretation. Notably, in the content analysis, the research concentrated on extracting manifest meaning (what has been said), the explicit or surface-level meaning from the data instead of latent meaning (what is intended to be said) or deeper meaning [25]. In extracting manifest meaning, the words in the text were used as they are, rather than finding the underlying meaning or deeper interpretation of the words or text [25].

Content analysis was chosen for the research as it is systematic, rigorous [276], and flexible for analysing text data using qualitative approaches [41; 248; 98]. In this research, content analysis facilitates evaluating the overall qualitative data analysis, improves rigour, and provides an understanding of data for constructing procedures, guidelines and practices [258]. Moreover, content analysis suits the exploratory nature of the research, best suited to generate findings and put them into the context of the research topic [276]. Four steps in the content analysis guided the research [25]; 1)Planning – define the aim, objectives, and research ethics reviews; 2) Data collection – choice of participants, data collection method, recording and transcribing the transcripts; 3) Data analysis – familiarising with data, coding process, generate categories, concepts or themes 4) Reporting and presenting result – summarise data into graphs, table and provide meaningful qualitative insights.

In the content analysis of this research, the data presented in Chapter 5 and Chapter 6 was analysed using coding method [195] to capture significant information, words, phrases in the data [195]. An in-vivo coding was used to capture exactly what had been said by the participants of the studies in Chapter 5 and Chapter 6. Hence, the codes derived from in-vivo coding are concrete and specific.

It should be noted that this research used two different analysis methods, thematic analysis and content analysis, to analyse semi-structured interview data in Chapters 4 and 6. This is due to the different contexts of the studies presented in these two chapters, which required different data analysis methods to answer the RQs.

Thematic analysis was adopted in Chapter 4 to identify themes or patterns across the interview data to answer RQ2. This is aligned with previous studies suggesting the suitability of thematic analysis to derive themes, trends and patterns in the interview data [33]. This is explained in Section 4.2.2. A content analysis was adopted in Chapter 6 to describe the interview data objectively and quantify specific concepts, words or phrases to answer RQ4 without looking for themes or patterns across the data. This is aligned with previous studies suggesting the suitability of content analysis when the focus is to describe, explain and categorise interview data objectively [203]. This is explained in Section 6.2.3.

**Coding.** The study used codes and coding to capture the emerging themes, concepts, ideas and categories underpinning the data and helping to organise and interpret the data [76; 195]. Codes are tags or labels assigned to raw data collected in a study, e.g., from interviews and focus groups, for analysis purposes [160]. Coding is an integral part of the qualitative data analysis process and allows data reduction and simplification for subsequent analysis [54]. Coding is essential in analysing qualitative data after the raw data is converted into a more organised format for easier inspection and understanding [202]. Coding aims to identify categories, themes, and messages that fulfil the research aim.

Extracting relevant codes from the data involved an iterative process by constantly moving back and forth between the whole data and the code extracts to ensure the data have been fully explored, interpreted and coded [61]. The iteration here meant repeating the process before generating the desired result. The goal is to get closer to the desired results with each repetition. The iteration also helped ensure important information from the data that is of interest to the study was not missed during coding.

Two primary approaches are used in qualitative data coding, inductive and deductive coding [1]. This research used the inductive coding approach instead of deductive coding. Inductive coding was used to identify meaningful codes from the data, letting codes emerge without pre-defined codes, contrary to deductive coding based on pre-defined codes [1; 33]. An inductive coding approach was preferred over deductive coding as the former helped to gain as much information as possible from the data without restricting it to pre-defined codes. Furthermore, the scarcity of literature on the research topic justifies the adoption of inductive coding. Inductive coding was used in both thematic analysis and content analysis in this research.

Within the broader inductive coding, two approaches were adopted for this research, descriptive and in-vivo coding [1]. Descriptive coding was used in the thematic analysis presented in Chapter 4. In-vivo coding was used in the content analysis presented in Chapters 5 and 6. Notably, descriptive and in-vivo coding approaches were adopted to align with the respective RQs.

In descriptive coding, the codes summarise and describe participants' views or opinions in a word or phrase preferred by the researcher without the researcher putting any interpretation on the codes [33]. Descriptive coding was used together with thematic analysis, which could be the initial step in thematic analysis. In descriptive coding, the data is described and organised systematically to allow identifying patterns or trends in the data [1]. Descriptive coding used in this research describes data at a surface level without any interpretation aligning with the "manifest" approach used in the thematic analysis. This aligns with the scope of a study that seeks to describe the data at the surface level to derive meaningful themes or patterns to answer the RQ.

In-vivo coding captures actual spoken words or other specialised words uttered by the participants and helps to better engage with participants in generating codes [22; 76]. An in-vivo coding was used in line with content analysis for this research to capture precisely the words or phrases used by the participants and put them into the study context. Hence, without using the researcher's words to describe the participant's words. An in-vivo coding ensures avoiding misinterpretations of what participants said and maintaining the context of the data [22]. This aligns with the scope of a study that intends to identify specific words, terms or phrases uttered by participants to answer the RQ.

It should be noted that the researcher coded all the data collected for this research herself. Therefore, there were no issues in establishing consistency or reliability of the coding process (in comparison to a situation in which several coders were involved, which would require further checks to ensure consistency and reliability). Although the study has addressed "between-coder" inconsistency by having one coder, the "within-coder" inconsistency may not have been addressed. To overcome this issue, the researcher reviewed the codes multiple times and iterated before coming up with a final list of codes. A limitation related to "within-coder" or "intra-coder" is explained in Section 8.3.

Chapters 4, 5, and 6 of the thesis will provide more chapter-specific coding processes used in this research. Chapter 4 describes "descriptive coding" while Chapters 5 and 6 describe "in-vivo coding".

### 1.4.3 Research Ethics Review

Generally, research ethics reviews are essential in any research involving human participants to protect participants, be transparent to participants, ensure the researchers are protected, and ensure accountability throughout a research [112].

It is a requirement by the University of Kent, United Kingdom, that all research, staff or student, funded or unfunded, that involves human participants, their tissue or data, non-human subjects, or other material ethical issues must undergo ethical review before initiation [256]. In line with the University of Kent's requirement, this research adhered to the United Kingdom's (UK) guidelines on research ethics reviews. One such guideline is the University of Kent's research ethics and governance to ensure the research activities were conducted with integrity and according to ethical, legal, and professional frameworks, obligations, and standards [256]. The Social Research Association of the UK promotes good practice in social research, including high standards of research ethics to protect research participants, ensure research is high quality, comply with legislation, reassure funders and maintain the excellent reputation of the sector [256].

All research activities reported in this thesis complied with the following sound practice principles, drawn from the Universities UK's Concordat to Support Research Integrity and the UK Research Integrity Office's Code of Practice for Research: excellence, honesty, rigour, integrity, cooperation, transparency and open communication, accountability, training and skills, safety, care and respect [256]. The ethical review process approved and recommended by the University of Kent Central Research Ethics & Governance Committee (CREAG) is proportionate, risk-based, and involves two stages [256]:

1. The first stage is the completion of a research ethics checklist. The checklist asks questions about the type of project to be undertaken and will identify, by a 'yes' response to any of the questions, those projects that require a further full review.

2. The second stage, full review for those projects that require it, involves completing a more comprehensive application form designed to gather all information about the research relevant for a full review.

The checklist, complete application form, and supporting documentation (project protocol, consent forms, participant information sheets, research instruments, recruitment advertisements, as appropriate) must be submitted to the relevant CREAG, and the researcher must receive final approval from the CREAG before initiating the research.

Four ethics approvals were granted by the University of Kent's Central Research Ethics Advisory Group (CREAG) to conduct this research involving human subjects. More chapter-specific details of the four ethics review approvals will be provided in Chapters 4, 5, and 6. Participants were provided with a Personal Information Sheet containing details of the studies and assurance to safeguard their personal data by complying with the UK General Data Protection Regulation (GDPR). The EU GDPR has been incorporated directly into UK law as the UK GDPR, which sits alongside the Data Protection Act 2018. A Consent Form was used to obtain informed consent to participate in the research without any inducements.

All participants willingly consented to participate in this research, recording their conversations and including direct quotations in research publications without disclosing their personal identifier information (PII). For privacy purposes, participants' identities were anonymised, and to provide context for specific quotes, pseudonymous identity documents (IDs) were used. Participants' affiliations were also anonymised as requested by participants. All personal data collected from the research are stored on the University of Kent's server, with access to the researcher through a Virtual Private Network (VPN) for security purposes. All data will be deleted securely and permanently from the University of Kent's server once the specified duration for keeping the data expires. It should be noted that no rewards or tokens were given to participants to show appreciation for their participation in the research.

More chapter-specific details of the overall methodology used in this research are explained in Chapters 3, 4, 5 and 6.

## 1.5 Research Original Contributions

The research reported in this thesis makes original contributions to knowledge and practice, as outlined below. The original contributions are further explained in Chapter 8.

1. Empirical evidence confirming that national CSIRTs use free tools and public data in responding to cyber security incidents. This contribution is towards answering RQ2.

2. Confirmation through empirical evidence that the use of free tools and public data by national CSIRTs is ad-hoc and not governed by formal, defined and institutionalised procedures. This contribution is towards answering RQ2.

3. A set of criteria to evaluate relevant attributes of free tools and public data. This contribution is towards answering RQ3.

4. Validation of the applicability of the aforementioned criteria in evaluating free tools and public data in national CSIRTs. This contribution is towards answering RQ4.

To support the validity of the research contributions – research publications, public presentations, external comments from peer-reviewed journals, and feedback from staff members of national CSIRTs during conferences have been considered. The preface of this thesis provides a list of relevant outputs from this research, which consists of published research papers and public presentations.

## 1.6 Thesis Organisation

This thesis contains eight chapters, with Chapter 1 already covered here. This section briefly explains the rest of the seven chapters to be covered in this thesis.

Chapter 2 provides some fundamental information about the research background to understand essential terms related to the research topic. The chapter defines security incidents in great detail, followed by explaining the definition of incident response, the importance of incident response, and how incident response fits into the overall cyber security field. The Computer Security Incident Response Team (CSIRT), the different types of CSIRTs, and their roles and responsibilities are described in this chapter. This is followed by a description of national CSIRT, a subset of CSIRT, and its roles and responsibilities in safeguarding the cyberspace of a nation.

Chapter 3 presents a literature review on the overall research topic identified from research papers, focusing on national CSIRTs' operations regarding free tools and public data. In addressing RQ1 of this research as defined in Section 1.3, this chapter also includes a section on a systematic literature review (SLR). In the SLR, a two-stage approach is adopted to gain systematic insights into the operational practices of national CSIRTs pertaining to free tools and public data. The research gaps identified in this chapter are revealed, and suggestions for future work to address these gaps are provided.

Chapter 4 presents and describes an empirical study using a survey and semi-structured interviews to understand the real-world operational practices of national CSIRTs regarding using free tools and public data to support incident response. This study addresses RQ2 of this research as stated in Section 1.3. The study was conducted in two phases. The first phase is to get some initial results of the operational practices from a single national CSIRT, followed by a second phase to enlarge findings from the first phase to gain more details of the operational practices across multiple national CSIRTs. Findings from the study revealed several gaps in key operational practices of national CSIRTs. One such gap is the lack of procedures to guide the systematic usage of tools and data. Hence, more research is suggested to systematically develop procedures for using tools and data.

Chapter 5 describes an empirical study using several focus groups to understand the current practices of free tools and public data evaluation practices across national CSIRTs. This study addresses RQ3 of this research as outlined in Section 1.3. Several gaps in current practices were identified; one is the lack of systematic evaluation of free tools and public data in national CSIRT. This informs the construction of a procedure for systematically evaluating free tools and public data. A set of candidate criteria for evaluating free tools and public data was constructed based on the opinions and views of staff members of national CSIRTs, presented in this chapter. Nevertheless, the validity of the candidate criteria remains uncertain. Hence, future work is recommended to validate the candidate criteria for usefulness, deployment readiness and applicability in practice.

Chapter 6 describes an empirical validation of candidate criteria for evaluating free tools and public data constructed in this research. This study addresses RQ4 of this research as defined in Section 1.3. The validation was conducted in 2-stage. First, validation to gain insights into how the staff members of national CSIRTs perceive the usefulness and deployment readiness of the candidate criteria in national CSIRTs. Second, a more objective validation by applying the criteria to evaluate two candidate tools and a data source for applicability in practice. The findings from the empirical validation of the candidate criteria are presented in this chapter.

Chapter 7 discusses the key findings of the overall research gained from the four significant studies reported in Chapters 3, 4, 5 and 6 of the thesis. This chapter discusses the key findings from the four aforementioned chapters in relation to the four RQs defined for this research.

Chapter 8 concludes the research by mapping the research works encompassing all four major studies reported in Chapters 3, 4, 5 and 6 to the four RQs, hence achieving the research aim. This is followed by summarising the research's original contributions. The research limitations and measures taken to mitigate the limitations are explained in this chapter. The mitigation prevents significant impacts on the overall validity of the research's findings. Finally, this chapter suggests several prospective future research that could expand the research reported in this thesis to the next level as a way forward to continue.

# Chapter 2

# Background

This chapter provides background information about the research topic to provide context to the research reported in this thesis. The background information includes support or evidence to strengthen the research topic and establish the topic's importance. Key concepts, terminology, historical information, relevant names, terms and organisations are introduced and described to provide a general understanding of the research. The chapter is structured as follows. Section 2.1 briefly explains security incidents, including definitions and what constitutes a security incident. Section 2.2 explains incident response, its meaning and importance. Section 2.3 describes the Computer Security Incident Response Team (CSIRT), which includes the different types of CSIRTs, roles, and responsibilities. Section 2.4 describes a subset of CSIRT, called the national CSIRT, its roles and significance. Section 2.5 explains the use of tools and data to support incident response in national CSIRT, and finally, Section 2.6 summarises the chapter.

## 2.1 Security Incidents

"*In January 2022, hackers attacked servers hosting the personal information of more than 500,000 people worldwide who receive services from the Red Cross and Red Crescent Movement. The hacked servers contained data related to the organisation's Restoring Family Links services, which reconnect people separated by war, migration, and violence.*" [245]. This is an example of a security incident, described in this section.

Any cyber attacks on the Internet affecting systems and networks are technically known as incidents or security incidents [161]. Other similar terms generally used in the literature and practice include computer security incidents and information security incidents. The ISO/IEC 27000:2018 uses the term information security incident, defined as a single or a series of unwanted or unexpected information security events that have a significant probability of compromising business operations and threatening information security [106]. Likewise, Ahmad et al. (2012) use the term information security incident to indicate when the confidentiality, integrity and availability (CIA) of an information asset is compromised [4]. Such information security incidents include malicious code, online fraud, distributed denial of service (DDoS), intrusion, intrusion attempts and data breaches [4].

The National Institute of Standards and Technology (NIST), an agency of the United States Department of Commerce, uses the term computer security incident [42]. It defined it as an occurrence that actually or potentially jeopardises the confidentiality, integrity, or availability of an information system or the information system processes, stores, or transmits or that constitutes a violation or imminent threat of violation of security policies, security procedures, or acceptable use policies [42]. Like Ahmad et al. (2012), the NIST pointed out that any event that compromised the CIA is a security incident. Cichonski et al. (2012) defined a computer security incident as a violation or imminent threat of violation of computer security policies, acceptable use policies, or standard security practices exemplified by the followings [42]:

1. An attacker commands a botnet to send high volumes of connection requests to a web server, causing the server to crash.

2. Users are tricked into opening a quarterly report sent via an email containing malware; executing the malware (unknowing it is malware) has infected their computers and established connections with an external malicious host.

3. An attacker obtains sensitive data and threatens that the details will be released publicly if the organisation does not pay a designated sum.

4. A user provides or exposes sensitive information to others through peer-to-peer file-sharing services.

The terms security incident and incident will be primarily used throughout the thesis. Other terms, such as cyber incident, cyber attack or cyber security attack, may be used occasionally to mean incident or security incident or computer security incident or information security incident. Notably, the term cyber threat denotes the probability of a cyber attack, while the term cyber crime refers to purely criminal activities.

Chapter 1 of this thesis informed that security incidents are increasing globally, devastatingly impacting organisations and citizens. Therefore, to deal with and handle the incidents, appropriate incident response is crucial and necessary in addressing incidents, as explained in the next section.

## 2.2 Incident Response

*The Red Cross system administrator took the affected servers offline to stop this attack [245]* – taking a server offline to stop an attack is an “incident response” described in this section.

Researchers have emphasised that security by prevention is not just enough; having appropriate response mechanisms to respond to security incidents is necessary [163]. Echoing Mitropoulos et al. (2006), Schneier (2014) pointed out that security is a combination of protection, detection, and response, with the 1990s being the era of protection while the current decade being the era of response [224]. This is exemplified by deploying a shield or protection consisting of policy and procedures, technology (e.g., firewalls, intrusion detection systems, antivirus software), physical controls, administrative controls (e.g., ISO/IEC Standards, NIST 800-53) and regulatory frameworks (e.g., GDPR, Payment Card Industry Data Security Standard (PCI-DSS)) as preventive measures against incidents [5]. Over time, the shield may not be able to prevent incidents as it may become vulnerable or pointless due to the sophistication of cyber-attacks that may potentially evade these "shields", causing potential cyber-attacks [5]. Therefore, ashield or protection alone is not sufficient.

This shows that incident response or computer security incident response is significant because it anchors the whole cyber security and information security field. It is the heart of information security. This means when an incident response is not deployed with appropriate response mechanisms, it may negatively impact the overall information security [163]. Therefore, incident response is critical to help minimise impacts and losses to organisations and information security at large [96]. Incident response is crucial to appropriately identify an incident, contain it, eradicate the root cause, restore it to a normal state and learn lessons from the incident [5]. Therefore, paying attention to incident response is equally critical, as experiences have shown a greater demand for effective handling of computer security incidents [133].

Incident response involves identification, containment, eradication, recovery, and lessons learned when examining a system that has been compromised [282]. It ensures continuity of operations during a crisis, mainly when a security incident occurs to organisations [282]. The advantage of incident response is that it supports systematic responses to incidents by following a well-defined procedure or plan so that the appropriate actions are taken in a timely and efficient manner [42]. The ISO/IEC 27035-1:2023 defines incident response as "actions taken to mitigate or resolve an information security incident, including those taken to protect and restore the normal operating conditions of an information system and the information stored in it" [107].

Incident response is a component of incident management, ensuring that controls and expenditures are entirely in place to detect, respond and mitigate cyberattacks [106]. It comprises a set of procedures for detecting, reporting, assessing, responding to, dealing with, and learning from information security incidents [106]. This procedure is in line with the requirement by the ISO/IEC 27002:2022 standard under "Clause 16.1.5 Response to information security incidents"[113]. The ISO/IEC 27035-2:2023 provides guidelines to plan and prepare for incident response and to learn lessons from incident response [108]. Notably, the Payment Card Industry (PCI) [44], and the Control Objectives for Information and Related Technologies (CobiT) [83] have established their incident response plans and procedures as essential components to respond to security incidents or breaches. This shows the importance of incident response to address and mitigate cyber attacks in various sectors.

In contrast, a lack of incident response may slow or delay the response to incidents or may even leave incidents not responded to. This may fail to comprehensively identify an incident, incomplete containment of the incident from spreading through the network, failure to eradicate the root cause of the incident, a complete restoration from an incident and application of preventive measures. Subsequently, this causes unwanted risks and impacts on information security in the long run, exemplified by revenue loss, damage to brand reputation or image and loss of customers' trust [228; 96]. Therefore, response to security incidents has always been important in the science of information security [163], requiring well-defined hierarchical procedures to respond to security incidents accordingly.

### 2.2.1 Incident Response Phases

Incident response is a linear and plan-driven process model consisting of sequential phases for handling incidents [42; 275] – prepare, identify, contain, eradicate, recover and follow-up (or lessons learnt). The ISO/IEC 27035:2023 standard provides guidelines to plan and prepare for incident response and to learn lessons from incident response [108]. The guidelines are based on the "plan and prepare" and "learn lessons" phases of the information security incident management phases model presented in ISO/IEC 27035-1:2023. The Incident Response Model by the National Institute of Standards and Technology (NIST 800-61 Model) [42], is an example of a well-accepted incident response procedure in the industry and government practices which incorporates the following distinct phases – Preparation, Detection & Analysis, Containment, Eradication & Recovery, and Post Incident Activity, as shown in Figure 2.

While the SANS Institute[1] , a private U.S. for-profit company founded in 1989 that specialises in information security, such as cyber security training and certifications, features six phases in incident response – preparation, identification, containment, eradication, recovery and lessons learned [139; 182]. Similarly, Mitropoulos et al. (2006) proposed more practical phases in responding to security incidents, synthesis of the NIST and the SANS Institute – preparation, identification, containment, eradication, recovery and follow-up [163], shown in Figure 3.

The six phases of incident response outlined by Mitropoulos et al. (2006) [163], Cichonski et al. (2012) [42], Stephen Northcutt (2001) [182] and West et al. (2003) [275]

[1] See `https://www.sans.org/`

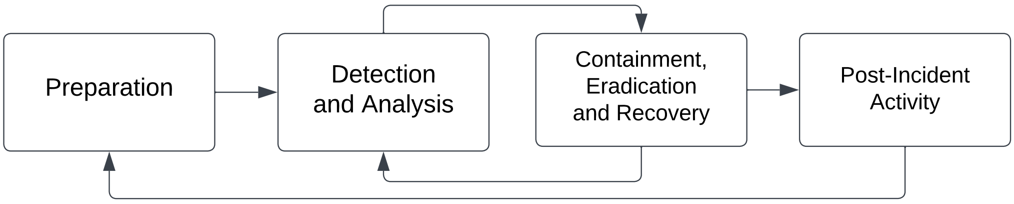


Figure 2: Incident Response Phases and Life-Cycle [42]

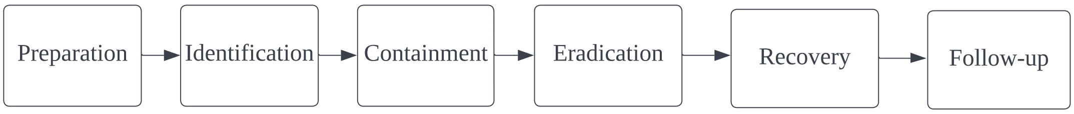


Figure 3: Incident Response Phases [163]

are described briefly below.

1. Preparation

   The objective of this phase is to be well prepared *before* an incident occurs. The concern in this phase is the "preparedness" and "readiness" to respond to an incident that may be unprecedented. In this phase, it is considered essential to establish a team, often called Computer Security Incident Response Team (CSIRT), which will be explained in the following section. Building relationships with key players of cyber security within a constituency, such as the government, public and private sectors, is crucial. In this phase, it is also critical to build response kits and other necessary tools and software deemed essential for the overall incident response [163]. Communication plans with internal and external parties and organising periodic cyber or tabletop exercises are essential to "readiness" before actual incidents occur. It is also vital to have plans to train staff with the necessary technical skills critical for incident response.

2. Identification

   Identification aims to identify an incident by gathering initial details and evidence from the potentially compromised systems. This includes assessing incoming incident reports, server logs, traces of incidents, and co-relating with other useful information [35]. This crucial step confirms an incident and moves on to the rest of the steps until the compromise is restored and preventive measures are considered [163]. The occurrence of an incident can be determined by first looking for simple mistakes made by the System Administrator. This might show a misconfiguration issue, eliminating the likelihood of a cyber attack. Evidence or artefacts related to incidents can be gathered at the network perimeters, i.e. firewall, router, intrusion detection system (IDS), intrusion prevention system (IPS) and demilitarised zone (DMZ) system [163]. Evidence can also be found at the host perimeter, such as personal firewall, IPS, local firewall, system-level antivirus (AV) tools, file integrity tools, endpoint security suite, and application level, such as application logs. Monitoring systems could be geared and enabled to help identify incidents [83]. Once an incident is confirmed, the following action will be to contain the incident or "containment", as explained below.

3. Containment

   The objective of Containment is to prevent an incident from further spreading to other networks and hosts or computers. This ensures the incident does not become widespread, making the response more complicated and time-consuming. Containment also helps prevent attackers from damaging an already compromised system further. One common containment practice is disconnecting the compromised system from the network [42]. Though isolating a compromised system from the network is a standard security practice, it depends on business requirements, as a certain amount of downtime will be involved. If an organisation under attack is planning to pursue legal action, this is the phase to preserve evidence without destroying or tampering with the evidence. Once the incident is successfully contained from further propagating, the next phase is eradicating and rectifying the incident's root cause and applying fixes, as explained below.

4. Eradication

   Eradication aims to eliminate an incident's root cause and remove incident residues or attackers' traces [42]. This is a critical phase to ensure the compromised system is completely clean from any traces of the attackers, such as backdoor programs and executable files or malicious Java scripts related to the malware. This ensures the compromised system is clean and safe before re-connecting online to resume business. This phase is also crucial and must not be overlooked, as it could cause further damage, incomplete rectification of the compromise, and the potential appearance of new incidents if thorough eradication is not done. In some situations, e.g. in a rootkit attack, reinstalling the operating system should opt to eradicate the incident. Once a compromised system is completely eradicated, the recovery process must follow through.

5. Recovery

   The objective of the recovery phase is to resume business and operations as normal as before the incident. This is accomplished by re-connecting the compromised system online after complete eradication. This involves restoring the system from a clean and good working backup to resume business. It is also crucial to verify the integrity of files in the system using a “file integrity utility’ such as MD5 or SHA Checksum and restore all binaries which could have been changed during an incident. It is also crucial to apply the latest patches and upgrades to the system to correct all known vulnerabilities. In this phase, it is advised to change passwords to more robust ones and remove unnecessary services. A quick security assessment of the system is recommended to ensure it is secure before reconnecting online. Once the restoration is completed, the system can be put online to resume business and move on to the follow-up phase described below.

6. Follow-up

   The follow-up objective is to “close” the incident appropriately, monitor the affected system, document the incident, and learn lessons from the incident for improvements when responding to future incidents. Notably, following up on the incidents with other related parties, such as CSIRTs, ISPs, and site owners, is also essential on the system’s status after re-connecting online [35]. This ensures appropriate incident closure and completes the incident response life cycle – the six phases. In this phase, meetings with Senior Management to sum up and de-brief the overall incident, assess the impacts, and review the strengths and weaknesses of current policy and procedure should be emphasised [163]. It is vital to improve the current policies and procedures from a technical and administrative aspect. Finally, a clear and concrete plan for implementing all these improvements must be established in this phase, which is critical when responding to future incidents.

This section explains that incident response is a step-by-step procedure for dealing with security incidents. It takes place under considerable time pressure in organisations with loads of information, information diversity and task uncertainty [239], which requires people, processes, and technology to understand and respond to the incident appropriately and effectively. Protection of systems is almost all technology-based, while detection requires equal proportions of people, processes, and technology. On the other hand, the response part is mostly people-based, with critical assistance from the process and technology, as one could not automate an incident. Incident response needs people because successful incident response requires thinking [224]. The “people” component of incident response, highlighted by Steinke et al. (2015) and Schneier (2014), is made up of what is known as a “computer security incident response team” (CSIRT) to collectively respond to security incidents in a team structure described in the following section.

## 2.3 Computer Security Incident Response Team

Schneier (2014) has pointed out that incident response needs people to perform the tasks in responding to incidents [224]. This claim is supported by a recent empirical study conducted as part of this research, with 17 national CSIRTs, which found manual approaches are predominantly used in 14 out of 17 national CSIRTs to accomplish incident responses [169]. This finding agrees with Schneier (2014), which shows incident response largely depends on manual approaches that need people, often in groups or teams, called an incident response team or Computer Security Incident Response Team (CSIRT).

The concept of CSIRT first emerged from team efforts and experiences handling the first ever "Internet worm" incident, which hit the Internet in November 1988, the Morris Worm attack [235]. This established the first CSIRT, the CERT/CC[2] , under Carnegie Mellon University, USA [233]. This happened after realising the urgent need for a specialised team to respond to incidents [271]. It should be noted that CERT/CC is now called the CERT Division of the Software Engineering Institute (SEI) of Carnegie Mellon University in the USA [233]. Since the establishment of the first CSIRT, CSIRTs have gained attention in the realm of cyber security for their significant role in responding to and mitigating security incidents [275; 163; 216]. Notably, many early CSIRTs were established by academic communities, for example, in the setup of CERT/CC of Carnegie Mellon University in 1988 [59] and the Australian Computer Emergency Response Team (AusCERT) [16], established in the University of Queensland, Australia in 1993.

The substantial need for incident response expanded the establishment of CSIRTs to various sectors worldwide – public and private. Over time, CSIRTs flourished worldwide, extending to the organisational, national (national CSIRTs) and regional levels. In the broader Internet community, CSIRTs form a "global network" from diverse organisations and sectors, such as critical infrastructure, government, industry, and academia [232]. Today, CSIRTs have been widely established within regions, countries, governments, the military, academics, products and businesses [9]. While national CSIRTs address issues on the national level, sector-specific CSIRTs address the cyber security needs of a specific sector, such as finance, health, transport, telecommunication, and utilities [110]. Other types of CSIRTs serve multinational companies, large companies, and private universities [110]. Organisational CSIRTs or Internal CSIRTs (sometimes referred to as "enterprise" CSIRTs) operate at the organisations' levels such as a private company, enterprise or businesses [59]. A list of different types of CSIRTs is summarised in Table 1. Cross-CSIRT Organisations is mentioned in Table 1 as it is closely related to CSIRTs. Notably, regional in Table 1 refers to a region that consists of multiple countries.

A CSIRT is responsible for receiving, reviewing, and responding to security incident reports and activities [88]. It consists of a "group of experts within an organisation which

[2] See `https://www.sei.cmu.edu/about/divisions/cert/index.cfm`

Table 1: Types of CSIRTs [225]

| **Types of CSIRTs** | **Constituency Served** |
|---|---|
| Regional CSIRTs | A region, e.g., Asia Pacific, Europe |
| National CSIRTs | A country |
| Public Sector CSIRTs | A part of a country, a locality |
| Governmental CSIRTs | National Government, State-level Government |
| Organisational CSIRTs | Any organisation |
| Product CSIRTs | Respond to product vulnerabilities, e.g., CISCO, Microsoft CSIRTs |
| Sector CSIRTs | A particular sector, e.g., Energy Sector CSIRT |
| Commercial CSIRTs | A commercial entity |
| Cross-CSIRT Organisations | NA |

is responsible for handling incidents related to IT security issues" [225]. The ISO/IEC 27000:2018 standard defined an "Information Security Incident Response Team" as appropriately skilled and trusted organisation members that handle information security incidents during the incident life cycle [106]. Apart from generally known as *computer security incident response team (CSIRT)* [216; 275], an incident response team is also called a *computer emergency response (or readiness) team (CERT)* [9], or a *cyber (or computer) incident response (or readiness) team (CIRT)* [109]. The ISO/IEC 27035-1:2023 uses the term *Incident Response Team (IRT)* defined as a team of appropriately skilled and trusted members of an organisation that responds to and resolves incidents in a coordinated way [107]. Notably, ISO/IEC 27035-1:2023 also clarified that Computer Emergency Response Team (CERT) and Computer Security Incident Response Team (CSIRT) are specific examples of IRTs in organisations and sectorial, regional, and national entities wanting to coordinate their response to large-scale ICT and cyber security incidents [107].

For this research purpose, this thesis uses the term "CSIRT" as it is commonly used in the research literature and by security practitioners.

ISO/IEC 27035-1:2023 recommends having a permanent incident response team (IRT) for respective constituencies [107]. Establishing a computer security incident response team (CSIRT) ensures a team trained and experienced in resolving incidents and coordinating with internal and external stakeholders on incidents [149]. Such a CSIRT responds typically to cyber incidents following an established *standard operating procedure (SOP)* [167]. This SOP is mainly adapted from the guidelines defined by standardisation in cyber security, such as the "Computer Security Incident Handling Guide" defined by the NIST in the US [42], the International Standard Organisation – ISO/IEC 27002:2013 [237] and the SANS Institute of the USA [139].

In principle, CSIRTs' services consist of reactive [271] such as incident management, incident response, and incident handling [216; 126], proactive such as monitoring, vulnerability handling [157; 97; 274] and information sharing within the team [186]. Several CSIRTs provide a combination of reactive (i.e., incident response and intrusion detection) and proactive services (i.e., vulnerability management, risk assessments, security consulting, and penetration testing) [267], while other CSIRTs focus on providing reactive services only. The high-level set of services a CSIRT might provide is summarised in Table 2. CSIRTs perform these services based on their mission, scope, expertise, and constituent requirements. In addition to the above services, CSIRTs participate in national and international research initiatives related to cyber security on identifying methods for detecting security incidents [184]. A full description of the work of CSIRTs can be found in RFC 2350, Internet Engineering Task Force, 1998 [35].

Table 2: Proactive and Reactive Services Provided by CSIRTs [216]

| **Reactive Services** | **Proactive Services** | **Security Quality Management Services** |
|---|---|---|
| Alerts and Warning | Announcement | Risk Analysis |
| Incident Handling | Technology Watch | Business continuity and disaster recovery planning |
| - Incident Analysis | Security audit or assessments | Security consulting |
| - Incident response on site | Configuration and maintenance of security tools, applications and infrastructures | Awareness building |
| - Incident response support | Development of security tools | Education and training |
| - Incident response coordination | Intrusion detection tools | Product evaluation or certification |
| Vulnerability Handling | Security-related information dissemination | |
| - Vulnerability analysis | | |
| - Vulnerability response | | |
| - Vulnerability response coordination | | |
| Artefact handling | | |
| - Artefact analysis | | |
| - Artefact response | | |
| - Artefact response coordination | | |

The literature made clear that CSIRTs play an outstanding role in the overall response to security incidents on which societies and national strategies rely [94]. Chapter 1 provided a comprehensive understanding of the current threat landscape, the frequency and seriousness of security incidents or incidents or cyber attacks. Therefore, the responsibility to respond to and mitigate the incidents is now in the arms of CSIRTs [252; 206], depending on the constituents' requirements. A much bigger responsibility is crucial at national levels to ensure that the entire cyber space of a nation is safeguarded. This responsibility often lies on national CSIRTs, a distinct type of CSIRT of particular interest for this research, described in detail in the next section.

## 2.4 National CSIRT

This section provides a general understanding of national CSIRTs' roles and activities in safeguarding a nation's cyberspace. National CSIRTs are also entrusted with coordinating with other CSIRTs within and outside their constituencies when handling large-scale and widespread incidents more effectively. This implies the importance of national CSIRTs for effective response and coordination of incidents at the national level, with practical implications for a nation and the world at large.

National CSIRT is a distinct category within the broader CSIRT landscape, and many economies worldwide have formed their national CSIRTs [239]. National CIRTs, or CSIRTs, allow countries to respond to and coordinate incidents at the national level through a centralised contact point more quickly and systematically, empowering countries to learn from experience and build cyber security resilience [110]. National CSIRTs are often developed and implemented following legislation or national policy. For instance, the Pan-European Cooperation Commission in 2009 requested its member states to ensure that national CSIRTs act as the vital component of national capability [63]. This includes preparedness, information sharing, coordination, response and leading national contingency planning and exercises [63].

In 2016, the European Parliament and the European Council passed the EU NIS Directive [64], which requires EU member states to have well-functioning national CSIRTs. This function is to "comply with essential requirements to guarantee effective and compatible capabilities to deal with incidents and risks and ensure efficient cooperation at Union level". In addition, Article 12 of the EU NIS Directive mandates cross-border collaboration between EU member states, including the pan-EU CSIRTs Network [36]. The United Nations (UN) body, the International Telecommunication Union (ITU), is actively promoting countries to establish national CSIRTs [109]. A list of countries with national CSIRTs from around the world is constantly compiled and available at the ITU [109]. At the end of 2020, 131 countries established national CSIRTs, including ten new CSIRTs. Since the 2018 Global Cybersecurity Index, four more national CSIRTs are currently under development [110]. As of 2020, Europe has the most significant number of national CSIRTs (41), followed by Asia-Pacific (26), America (21), and Africa (19) [110].

National CSIRTs may have various focus areas or constituencies and mandates depending on the requirements of the country in which they are situated. National CSIRTs can be part of a governmental institution or under the umbrella of a specific ministry or entity [110]. A national CSIRT can also be situated outside of governments, but it must have a national responsibility, recognised by a government in the country or economy [27]. In some countries, national CSIRTs may be positioned under a national authority or agency in charge of cyber security matters [60], which has a broader mandate as a national cyber security centre. Table 3 shows examples of how national CSIRTs are embedded differently in their respective countries worldwide.

Table 3: How National CSIRTs are Embedded in Different Ways Worldwide [60]

| **National CSIRT institutional embedding as part of:** | **Example of National CSIRTs** |
|---|---|
| Prime Minister's Office | CERT VU (Vanuatu), CERT-BE (Belgium) |
| An agency under the supervision of a Ministry (Interior, ICT, Environment, etc.) | ThaiCERT (Thailand), CERT-GH (Ghana), CERT Tonga (Tonga) |
| Communications Regulatory Authority | TZ-CERT (Tanzania), NCSC-FI (Finland), CARICERT (Curacao) |
| National security authorities | TTCSIRT (Trinidad Tobago) |
| National Defence<br>Cyber Security Agency | Hellenic CSIRT (Greece)<br>SingCERT (Singapore), CERT-SA (Saudi Arabia) |
| National Cyber Security Centre (NCSC) | NCSC (New Zealand), Canadian Centre for Cyber Security (Canada), NCSC-NL (The Netherlands) |
| Domain name registrar | CERT.br (Brazil), CERT.at (Austria) |
| Private limited liability | Sri Lanka CERT\|CC (Sri Lanka), BruCERT (Brunei) |

A country may have a single national CSIRT or more than one, depending on its legal needs and requirements. Their roles and functions can be shared between more than one team [184]. For instance, the Polish legislators have designated three national-level CSIRTs for Poland: the CSIRT Ministry of National Defence (MON), CSIRT of National Research Institute (NASK), and CSIRT of Internal Security Agency (GOV) [184], whereby all three national CSIRTs work together to coordinate and respond to incidents affecting the Poland constituency. To avoid competition among the three national CSIRTs in Poland, the Legislators define the roles of each of the national CSIRTs [184]. Similarly, Switzerland has two national CSIRTs, the SWITCH-CERT and the NCSC-CH. The former provides services to research and education, the domain registry, banks, industry, logistics, and the energy sector and the latter to the government sector.

A national CSIRT's key role is safeguarding a nation from cyber attacks and as a national point of contact for matters related to cyber attacks [164; 91; 173]. This includes to respond and coordinate responses to cyber attacks [279; 9; 60; 214], and protecting critical national infrastructures from cyber attacks [94; 186; 9]. National CSIRTs also support citizens and organisations experiencing cyber attacks on remediations [186], sharing indicators of compromise with other national CSIRTs and alerting the constituency of potential cyber threats [167]. Their role is similar to fire brigade emergency services, consisting of reactive and proactive services to organisations and citizens in a nation that are not commercial and business oriented [275; 226]. Overall, national CSIRTs' roles can be summarised as followings [27]:

1. Being a national focal point within a country, economy, or region to coordinate incident handling activities, which includes cooperation with the global National CSIRT community.
2. Analyzing and synthesizing incident, vulnerability, and threat information disseminated by their constituency, other teams, vendors, and technology experts to provide an assessment for their constituency and communities.
3. Facilitating communications across a diverse constituency – bringing together multiple sectors (e.g., government and military, critical services and infrastructures, commercial, academic, banking and finance, transportation) to share information and address computer security problems, such as widespread computer security incidents, threats, and vulnerabilities.
4. Developing mechanisms for trusted communications within the communities.

Besides the above roles, some national CSIRTs also engage in the following activities [110]:

1. Promoting cyber security awareness and online child protection by providing tips, guides, manuals, training, and videos
2. Delivering cyber security advisories to IT specialists
3. Conducting cyber drills or cyber exercises
4. Engaging with regional CSIRTs and FIRST
5. Getting certified by Trusted Introducer or other recognised certification.

## 2.5 Understanding Tools and Data

Tools and data are essential armours to help facilitate incident response and mitigating incidents in any operations of CSIRTs, including in national CSIRTs. Hence, the following section describes tools and data and why they are relevant and vital to support incident responses in national CSIRTs.

National CSIRTs typically use tools and data from various sources to support incident responses. Such tools and data are paramount for national CSIRTs and CSIRTs in general *to perform effective and efficient incident responses* [111; 27]. Having *good and effective* *mitigation tools for CSIRTs* and companies at large are essential in mitigating cyber attacks [101]. Equally, it is important to use good-quality tools and data to facilitate incident responses in national CSIRTs, which can be achieved through evaluating and implementing tools [27]. Using appropriate tools to facilitate incident responses can significantly increase the effectiveness of CSIRTs.

The definitions of tools and data are not consistent across the literature. The Cambridge Dictionary[3] defines a tool as *a program or feature of a program that helps you do particular things on a computer: The spreadsheet program offers several useful tools for manipulating the material. Companies in the industry tend to use identical, or at least compatible, software tools.* The Cambridge Dictionary defines data as *information, especially facts or numbers, collected to be examined and considered and used to help decision-making or information in an electronic form that can be stored and used by a computer: the data was/were collected by various researchers. Now the data is/are being transferred from magnetic tape to hard disk.*

The Merriam-Webster Dictionary[4] defines a tool as *an element of a computer program (such as a graphics application) that activates and controls a particular function.* The Merriam-Webster Dictionary defines data as *1) factual information (such as measurements or statistics) used as a basis for reasoning, discussion, or calculation and 2) information in digital form that can be transmitted or processed.*

There is no precise definition of tools in the ISO/IEC Standard, except for a definition of software products. The ISO/IEC 25010:2011 [104] defines software products as a *set of computer programs, procedures, and possibly associated documentation and data.* Data is defined as *reinterpretable representation of information in a formalised manner suitable for communication, interpretation, or processing* by the ISO/IEC 25012:2008 [103].

Grundy et al. (2001) defined tools as *Tools are used to describe and share software processes, plan and manage teamwork, capture process performance information and ultimately improve processes* [84]. The authors also differentiated different tools as having different functions; for example, *Computer-aided specification and design tools support software analysis and design and often incorporate round-trip engineering features, including code generation and reverse engineering.* While there is some information about the definition of tools, only a little information is available on the explicit definition of data. Nevertheless, it can be understood that data is information and knowledge which supports decision-making [194].

In general, tools can be classified into the below major types:

1. Free tools – refers to software tools and online services available to anyone for free. A major class of free tools is open-source tools, which are free tools whose developers publish the source code for anyone to see.

[3] See `https://dictionary.cambridge.org/`

[4] See `https://www.merriam-webster.com/`

2. Free online services – are considered free tools because, semantically, they are just software tools running from a remote server.

3. Commercial tools – are those that one has to pay for access to.

Tools are used to facilitate incident responses at each phase of incident response. Typically, different tools are used at each phase to facilitate incident response, though, in some instances, a similar tool can be used at various phases. For example, during the identification and eradication phases, most national CSIRTs use popular open-source tools such as Cuckoo[5] , Virus Total[6] and Malheur[7] to identify and eradicate malware. In the preparation phase, most national CSIRTs use open-source ticketing systems such as RTIR and RT for incident management [169].

In principle, data can be classified into the following major types:

1. Public data – refers to data available to the general public, often freely but sometimes behind a paywall (in the latter case, there is no restriction on data consumers, i.e., anyone can pay and access the data). CSIRTs use public data, such as newsletters and advisories from other national CSIRTs, to support their operations.

2. Open data – different from "public data" based on the attributes of the data source and structuredness. Open data "can be freely used, modified, and shared by anyone for any purpose" [265]. Open Data is an emerging term in defining how scientific data may be published and re-used without price or permission barriers and open to everyone for free [175].

3. Closed-source data – refers to data unavailable to the general public but to selected data consumers only (e.g., CSIRTs, public bodies, and people with a security clearance status). This includes self-reported incident data or evidence from victims (including organisations and citizens), law enforcement agencies (LEAs), and trusted partners such as collaborating CSIRTs [165; 166].

## 2.6 Summary

This chapter presented comprehensive background information about the research and the fundamental knowledge required to undertake the research. The chapter introduced several key concepts related to the research (incident, incident response, CSIRT, national CSIRT) and described how these concepts relate to one another coherently. The definitions of tools and data and their usefulness for incident response and national CSIRTs

[5] See `https://cuckoosandbox.org/`

[6] See `https://www.virustotal.com/`

[7] See `https://github.com/rieck/malheur`

were laid out to provide context for this research. Nonetheless, many unknowns exist about national CSIRTs' operational practices concerning tools and data to facilitate incident response. More specifically, the use of free tools and public data in national CSIRTs is less understood and studied. There is a need for an in-depth literature review to gain a better understanding of the operational practices of national CSIRTs' free tools and public data. Systematic information about national CSIRTs can help substantiate the current understanding and systematically identify gaps and areas for improvement in national CSIRTs. In the next chapter, a systematic understanding from the literature of the operational practices at national CSIRTs concerning free tools and public data to facilitate incident response is laid out.

# Chapter 3

# Literature Review

This chapter provides an overview of the literature (LR) on the overall research topic, focusing on operational practices concerning tools and data at national CSIRTs. A systematic literature review (SLR) to gain systematic insights specifically into the operational practices of national CSIRTs pertaining to free tools and public data is also presented in this chapter. While the LR searches for scientific sources or papers, the SLR systematically searches the websites of national CSIRTs and cross-CSIRTs and the scientific database to gain insights into the operational practices of national CSIRTs concerning free tools and public data. This chapter analyses the current operational practices of national CSIRTs, focusing on the existing operational strengths and shortcomings. This allows for identifying research gaps, which can be used as the basis for this research. After reading this chapter, a better and more systematic understanding of the operational practices of national CSIRTs should be achieved. This includes the gaps and areas for enhancing the current operational practices. Eventually, this leads to the remainder of the thesis addressing the identified gaps.

The chapter is structured as follows. Sections 3.1–3.4 reviews the literature related to the overall research topic of this thesis with a focus on operational practices at national CSIRTs concerning tools and data. A systematic literature review (SLR) to gain systematic insights specifically into the operational practices of national CSIRTs pertaining to free tools and public data is presented in Section 3.5. The research gaps identified from this chapter are presented in Section 3.6. Finally, Section 3.7 summarises the chapter.

## 3.1 National CSIRTs' Operational Practices

According to Werlinger et al. (2010), incident response is a highly collaborative work, which very often requires practitioners to develop in-house tools for diagnostics and detection of incident tasks [274]. Furthermore, detecting incidents is complicated, requiring expertise on the part of the staff and adequate support from usable and reliable security tools. Even though Werlinger et al. (2010) did not study national CSIRTs' operations, it provides a general understanding of incident response practices and the common problems encountered in real-world CSIRTs' operations. Moreover, it is crucial to have usable and reliable tooling to support incident responses [274]. Besides, tools to support efficient information sharing and communication are also vital for incident response and help improve current CSIRT capacity [115].

Elaborating on information sharing, threat data and other threat information needed for incident response are actively shared among internal CSIRTs within the Eurocontrol Air Traffic Management CERT (EUROCONTROL/EATM-CERT) and with several national CSIRTs in Europe [154]. Threat information within incident response practices is shared primarily via tools such as the Malware Information Sharing Platform (MISP)[8] . Besides information-sharing and effective threat detection tools, honeynets are equally essential to support incident responses by better understanding current cyber threats and their modus operandi [125]. Automation is much needed for more effective incident responses [250]. Tondel et al. (2014) highlighted the importance of tools support for incident response and automation in the operational practices of CSIRTs. In addition to tools supporting incident response, classified and closed-source data are essential to responding to and detecting threats more efficiently. Grispos et al. (2015) studied the importance of data quality needed for post-incident learning [82]. This points to the diverse need for quality data to support incident responses.

All the above studies from the literature are about the operational practices of CSIRTs, primarily in organisations and studies about national CSIRTS are very scarce. Very little work was done on understanding real-world operational practices of national CSIRTs concerning free tools and public data in facilitating incident responses and how the staff perceive such tools and data. The topic of CSIRTs, in general, is mainly under-represented in academic research due to the novelty of the topic itself [140], and what's more, on national CSIRTs' operations. This suggests the need for more systematic and empirical studies about national CSIRTs to gain more insights by exploring the many areas within national CSIRTs that are currently less understood [149]. Doing so contributes to enhancing the current knowledge and provides foundations for future research on national CSIRTs [186].

## 3.2 Importance of Quality Tool and Data for National CSIRTs

Tools and data are paramount for CSIRTs and national CSIRTs *to perform successful and effective incident responses* [111; 27]. Having *good and effective mitigation tools for*

[8] See `https://www.misp-project.org/`

*CSIRTs* and companies at large are essential in mitigating cyber attacks [101]. Using appropriate technological tools to facilitate incident response can significantly increase the effectiveness of CSIRTs. The lack of efficient tools could cause issues in cyber incident responses, leading to greater security risks [58] as using effective and qualified tools in CSIRTs' operations is vital to prevent and mitigate cyber attacks [101]. The International Telecommunications Union (ITU) has highlighted the importance of CSIRTs and how the establishment and development of CSIRTs should be based on mature models, utilising international collaborations, robust procedures, *effective tools* and training [58]. Effective tools may reside in the lead time of solving an incident, on the financial level and in increasing team knowledge and shared situation awareness within a CSIRT [131].

Several researchers emphasised that having qualified tools and data is crucial in responding to cyber-attacks and incidents more efficiently [207; 71]. The need for quality tools and data is increasing as more and more organisations rely on tools and data in many aspects of their operations [243]. These include national CSIRTs, extensively using software tools (in particular, open-source and free tools), public data, and open-source intelligence (OSINT) to facilitate incident responses [167]. Hence, tools and data must be evaluated systematically by following a specific set of criteria for quality purposes [87]. Such evaluation ensures compliance with security requirements [28] and effectiveness for operations in national CSIRTs [27].

Furthermore, to ensure only qualified tools and data are selected, particularly for incident management and analysis work in national CSIRTs, the evaluation and implementation of tools is necessary [27]. CSIRTS must examine IT devices or software to identify vulnerabilities; such examination is crucial to avoid using unpatched and vulnerable tools in the operations [184]. Therefore, evaluating tools and data in national CSIRTs and CSIRTs at large is essential to ensure the team is equipped with quality tools and data to detect, respond, and mitigate security incidents effectively.

All the above studies highlighted the importance of effective tools and data to support national CSIRTs' operations, giving the impression that qualified, free tools and public data are essential to facilitate effective incident response. Identifying such tools and data can only be achieved through systematic procedures for evaluating tools and data using a set of criteria that have been validated for usefulness, deployment and applicability in practice. Hence, this study intends to establish criteria for evaluating tools and data for national CSIRTs that could benefit broader security operations.

## 3.3 Evaluation of Tools and Data to Ensure Quality

**Tool evaluation.** Studies and past research concerning tool and data evaluation within CSIRTs or national CSIRTs are sparse in the literature. Though researchers have investigated the general use of tools in the operational practices of CSIRTs [140; 236], a vital area of research – systematic procedures to evaluate tools and data is missing. No attempt has been made to study the evaluation of free tools and public data in national CSIRTs. Such research ensures qualified tools and data support incident responses in national CSIRTs.

A need-assessment study using interviews with CSIRTs identified certain features (criteria) that must be present in tools to indicate tool efficiency [131]. These include the ability to produce reports and output that are more structured and reader-friendly, the scalability of tools to handle large-scale incidents, visualisations for a better understanding of the insights, and the ability to support different levels of details. Echoing Kleij et al. (2017), this research is interested in identifying features or criteria that can be used as a baseline to select effective tools. Enlightening on the need for tool evaluation, Iakovakis et al. (2021) developed a list of mitigation tools by the following criteria – strengths, weaknesses, free trial, cost/price, scalability, technical support, vulnerability assessment, reports and analytics, ease of use, GUI offered, and compatibility [101]. However, methodologically, it is unclear how the authors identified such criteria. The research reported in this thesis intends to extend the works done in previous studies to develop a procedure consisting of criteria that can be used for evaluating tools and data for quality purposes, with an explicit methodological process.

**Data evaluation.** Problems with data quality have been highlighted in previous studies, such as a lack of metadata and incomplete data [265]. Some of these problems have been addressed by proposing a new design of data quality management (DQM) for open data based on the ISO/IEC 25010 and ISO/IEC 25012 for data quality management (DQM) [265]. The authors proposed a new DQM framework for open data usable in the agricultural sector. They envisaged that the results could be further exploited to develop an intelligent system with open data aggregation.

Notably, some work has been done in evaluating data within CSIRTs, more precisely, on threat data feeds. This includes evaluating threat intelligence data feeds to help users choose qualified threat data feeds available on the Internet [201]. The authors evaluated sample threat feeds using the following factors (criteria): relevance, accuracy, completeness, timeliness and ingestibility, using several metrics to identify the most qualified feeds. Similarly, a study evaluated blocklist data feeds based on the following criteria: vantage, volume, timeliness, accuracy, and completeness using several metrics [142].

It should be noted that the authors used a set of criteria adapted from the literature in the above studies. In contrast, the research reported in this thesis intends to develop new criteria for evaluating public data through the collective opinions of national CSIRT staff members. This aligns with the call for organisations to have ongoing processes to evaluate data such as vulnerability information by formalising such evaluations with suitable metrics and measurements [52]. Doing so helps security practitioners to choose quality data and eliminate problematic data such as incomplete and untimely data for incident response. Notably, the above studies also emphasised the need to test the factors identified in the studies on the applicability for the respective environment [247]. Similarly, such testing should apply to tool evaluation criteria, too.

A data quality evaluation environment was outlined based on the international standards ISO/IEC 25012 (which defines the data quality characteristics) consisting of accuracy, completeness, currentness, consistency and credibility that can be used to evaluate data quality [85]. Similarly, another study developed a data cyber security model which defines data quality characteristics selected from the ISO/IEC 25012 standard (data quality model) consisting of compliance, confidentiality, traceability, availability and recoverability [263]. The model was applied to an existing commercial product to validate the framework and verify the feasibility of this framework for evaluating data.

## 3.4 Understanding Tool and Data Evaluation Practices

### 3.4.1 Tool and Data Evaluation Models

At the core of software tools is software evaluation and quality assessment. In the literature, several software evaluation models have been mentioned and proposed [10; 229; 50]. The four commonly mentioned ones are: McCall Model (1977) [156], Boehm Model (1978) [28], FURPS Model (1992) [81] and Dromey Model (1995) [57].

The McCall Model introduced in 1977 contains a set of eleven factors (derived from a literature survey) for evaluating software, specifically for the US Air Force applications [156]. Fifty-five factors associated with software quality were identified from the literature and grouped into eleven software quality factors: *maintainability*, *flexibility*, *testability*, *portability*, *reusability*, *interoperability*, *correctness*, *reliability*, *efficiency*, *integrity* and *usability*. These factors were then validated with two US Air Force systems. The results were translated into guidelines that Air Force Program Offices could use to specify software quality. The downside of the McCall Model is that the functionality and compatibility of software products were not directly considered. Security aspects of software were not considered, either.

The Boehm Model was introduced in 1978, which consists of a hierarchical quality model with the following factors: *portability*, *reliability*, *efficiency*, *testability*, *understandability*, *human engineering* and *modifiability* [28]. The Boehm Model addressed the shortcomings of other software evaluation models that focus on precisely measuring high-level characteristics. Hence, Boehm introduced three new factors to the McCall Model: *understandability*, *human engineering* and *modifiability*. However, the Boehm Model fails to address the functionality, compatibility and security factors for evaluating software.

Robert Grady and Hewlett-Packard Company proposed the FURPS Model in 1992 [81]. This model classifies software quality characteristics into two categories: "functional" and "non-functional" requirements. The input and the expected output define functional requirements. Non-functional requirements defined by *usability*, *reliability*, *performance*, *supportability.* The IBM Rational Software extended the model into FURPS+, where the "+" indicates design constraints, implementation, interface, and physical requirements. However, the models fail to consider product characteristics such as *portability*, *understandability*, *flexibility* and *learnability*, making the model incomplete.

Geoff Dromey designed the Dromey Model, a product-based quality model, in 1995 [57]. Dromey believed quality evaluation is different for each product. Hence, a more dynamic view is needed to develop a model that is broad enough for various systems or products. The Dromey Model focuses on understanding the relationship between the attributes (characteristics) and the sub-attributes (sub-characteristics) of quality in software. The problem with the Dromey Model is that it does not address security and compatibility when evaluating software. We view security and compatibility as essential to align with current requirements for secure and dynamic tools.

### 3.4.2 Relevant Standards on Tool and Data Evaluation

In addition to the ISO/IEC 25012 standard mentioned in the previous subsection, two other international standards have been defined for software evaluation. The following three most relevant international standards were reviewed [159]:

1. ISO/IEC 25010:2011 Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – System and software quality models [104]

2. ISO 9241-11:2018 Ergonomics of human-system interaction – Part 11: Usability: Definitions and concepts [105]

3. ISO/IEC 25012:2008 Software engineering – Software product Quality Requirements and Evaluation (SQuaRE) – Data quality model [103]

The first two standards describe principles for assessing software quality [26] and the third one for data quality [269]. These standards have been widely adopted by researchers for software evaluation [178; 285; 253; 177; 135; 251; 12] and data evaluation [269; 283].

Notably, ISO/IEC 25010:2011 and ISO 9241-11:2018 are more recent and comprehensive than the McCall, Boehm, FURPS and Dromey Models for software quality evaluation. For instance, essential factors or criteria such as *security*, *satisfaction*, *functionality* and *learnability*, lacking in the McCall, Boehm, FURPS and Dromey Models, are addressed in the ISO/IEC 25010:2011.

### 3.4.3 Tool and Data Evaluation from User Perspective Context

The importance of including end users in software evaluation has been pointed out in a previous study [272]. This includes assessing the quality of the tools from the viewpoints of producers, operators, users, managers, and stakeholders using a comprehensive tool evaluation framework [29]. Users' perspective could help identify a tool's most and least satisfactory attributes, for instance, by using surveys to get users' opinions [17]. Previous studies were conducted to evaluate sample tools from users' perspective. Gade et al. (2013) used an online survey to evaluate sample tools with users and discussed the results in a focus group discussion [73]. The factors used in the evaluation were adopted from the ISO/IEC 9126-1:2001 Model [102]. Botchway et al. (2021) adopted factors from the ISO/IEC 9126-1:2001 Model and McCall's, Boehm's, FURPS and Dromey's Models to evaluate tools from users' perspective [30]. Stojkovski et al. (2021) used the User Experience Questionnaire (UEQ), a validated instrument for measuring User Experience (UX) on using software [242].

Notably, the majority of the user-based software and data evaluation studies from the literature adopted criteria or factors from ISO/IEC standards and software evaluation models, e.g., the ISO/IEC 9126:2001 Model, McCall, Boehm, FURPS, Dromey. To my knowledge, minimal studies have attempted to construct new factors or criteria from users' opinions. Hence, it would be interesting to explore criteria from recruited participants' views for inclusion to build a final set of criteria for evaluating tools and data.

## 3.5 Systematic Literature Review of National CSIRTs' Operational Practices

This section provides a Systematic Literature Review (SLR) to gain insights into the operational practices of national CSIRTs regarding the use of free tools and public data. The study was undertaken to answer RQ1 of this research defined in Section 1.3. The SLR systematically searches the websites of national CSIRTs, cross-CSIRT organisations, and the scientific database to identify relevant data items. This is followed by systematically synthesising and analysing the data items into themes that answer the RQ.

### 3.5.1 Introduction

The current threat landscape with increasing cyber attacks demands organisations to be more "cyber-resilient" with more effective approaches to defending against cyber attacks [246]. Although the system and network defence are essential, responding to attacks in a timely and efficient manner – quite often, facilitated by a dedicated incident response team [99] is equally important [43]. This includes having effective incident response teams with experience, technical skills, and capacity to respond to cyber incidents [63]. Such a team is essential to help respond to incidents and lower organisational risks due to incidents [216].

It is now known that tools and data are at the core of the operations of national CSIRTs. Many of the tools and data used in national CSIRTs are almost free tools and public data [167]. This gives the impression that free tools and public data are widely used in national CSIRTs. However, their use has something that is less understood and studied. The aim of the SLR is to understand the operational practices of national CSIRTs concerning free tools and public data  and *identify areas for future research that could improve current operations at national CSIRTs.*

The findings from the study can be summarised as follows:

1. A comprehensive systematic review of the literature on the operational practices of national CSIRTs concerning the use of free tools and public data to better understand the operations of national CSIRTs.

2. A list of 24 research papers identified through a systematic literature search concerning the use of free tools and public data in the operational practices of CSIRTs and national CSIRTs. The 24 research papers identified from the SLR help to advance research in this area.

3. The SLR identified research gaps that contribute to informing potential research and developments on national CSIRTs' operations.

4. A list of 299 free and open source tools, identified through a systematic search into national CSIRTs' websites, Cross-CSIRT organisations and research publications. This list is made available as an open resource for national CSIRTs, CSIRTs and security practitioners for use in real-world security practices at (`https://cyber.kent.ac.uk/research/CSIRTs/SLR/List_Open-source_Free_Tools.xlsx`). The list provides more open resources on free and open-source tools that practitioners can refer to.

The rest of this Section is organised as follows. Section 3.5.2 describes the methodology adopted for the SLR, and the results are described in Section 3.5.3.

### 3.5.2 Methodology

A Systematic Literature Review (SLR) was preferred for the study over a Literature Review (LR) because the former is a more rigorous and reproducible method for conducting literature reviews [129]. Traditionally, data items covered in SLR are research papers identified through systematic searches into scientific databases. However, to answer the RQ of this research, it is insufficient to consider just research papers because operational practices of national CSIRTs are more often published in other forms of non-research publications (e.g., operational documents, guidelines, articles on websites, training manuals, and even informal online discussions) [236]. Furthermore, CSIRT practices generally cover professional and business aspects; therefore, the sources on CSIRTs should not be limited to research literature only [236].

The SLR was conducted following a two-staged approach, Stage 1 and Stage 2:

- Stage 1a: A systematic search to identify relevant information (web pages and online documents such as technical reports, training manuals, articles, best practices and guidelines) accessible on websites of national CSIRTs and cross-CSIRT organisations.
- Stage 1b: A systematic search to identify relevant research papers in the Scopus Database.
- Stage 2: A synthesis of data items identified from websites of national CSIRTs, including cross-CSIRT organisations and research papers of the Scopus Database, to answer the research question.

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) procedure [148; 129; 95], a preferred SLR procedure used in many research fields, was adopted for the SLR. The PRISMA procedure consists of four main phases as below:

- Phase 1: Identification of data sources and candidate data items.
- Phase 2: Screening candidate data items according to exclusion criteria (based on checking metadata).
- Phase 3: Checking the eligibility of candidate data items according to inclusion criteria (based on reading full-text).
- Phase 4: Inclusion of the data items according to inclusion criteria.

As the searching and identifying of the relevant data items for the SLR was conducted in Stage 1a and Stage 1b, the PRISMA procedure was adopted independently, as explained in the below paragraphs.

**Stage 1a.** The websites of national CSIRTs and cross-CSIRT organisations were identified as the data source of non-scientific papers – *PRISMA Phase 1.* Next, the websites are screened, followed by screening the data items according to exclusion criteria (by checking the metadata) – *PRISMA Phase 2.* Then, check the eligibility of the candidate data items according to the inclusion criteria by reading the full text of all the screened data items – *PRISMA Phase 3.* Finally, after confirming the eligibility according to inclusion criteria, the data items are included for review and analysis – *PRISMA Phase 4.*

**Stage 1b.** The Scopus database was identified as the data source of research papers – *PRISMA Phase 1.* Next, the research papers were screened according to exclusion criteria (by checking the metadata) – *PRISMA Phase 2.* Then, check the eligibility of candidate research papers according to inclusion criteria by reading the full text of all the screened research papers – *PRISMA Phase 3.* Finally, after confirming the eligibility according to the inclusion criteria, the research papers are included for review and analysis – *PRISMA Phase 4.*

The scope of free tools in this SLR is limited to those designed for and used by national CSIRTs to facilitate incident responses. Hence, the following major categories of tools that are less relevant are excluded:

1. Tools designed for citizens and non-CSIRT organisations (e.g., malware clean-up tools, ransomware decryptors, cyber security awareness and educational tools).

2. Tools for purposes not directly related to incident responses (e.g., programming tools and standard OS commands, administrative tools, internal information management, human resource and finance systems).

3. Tools primarily used by cybercriminals, even if CSIRTs sometimes use them for security evaluation and research purposes. This is because there are too many such tools (e.g., every single piece of malware can be considered as such a tool).

4. Tools that do not have a software component (e.g., procedures, conceptual frameworks, recommendations and standards, guidelines and policies, pure hardware solutions).

5. Tools mentioned on a national CSIRT's web page without explicit evidence that a specific CSIRT used or is using it.

Additionally, shareware and commercial tools with a free option (with limited features or quota) were included, and those that expire after a defined trial period were excluded. The above criteria were applied to Stages 1a and 1b to identify relevant data items.

The different aspects of the procedure are explained in great detail. First, the procedure for searching the websites of national CSIRTs and cross-CSIRT organisations – in Stage 1a. This is followed by the procedure for searching the research papers – in Stage 1b.

### Stage 1a – Identifying Relevant Data Items on Websites of National CSIRTs and Cross-CSIRT Organisations

The procedure for Stage 1a is shown in Figure 4. The $w$ represents the websites of national CSIRTs and cross-CSIRT organisations, and $n$ represents the number of data items – the web pages and online documents. This stage is further split into three steps: 1) identifying and screening the data sources (the websites of national CSIRTs and cross-CSIRT organisations)–($w$), 2) searching into the websites (data sources) using the selected keywords to identify relevant data items–($n$) and 3) reading in full the data items to identify only relevant data items–($n$). The search in Stage 1 was conducted between 19 and 22 August 2021.

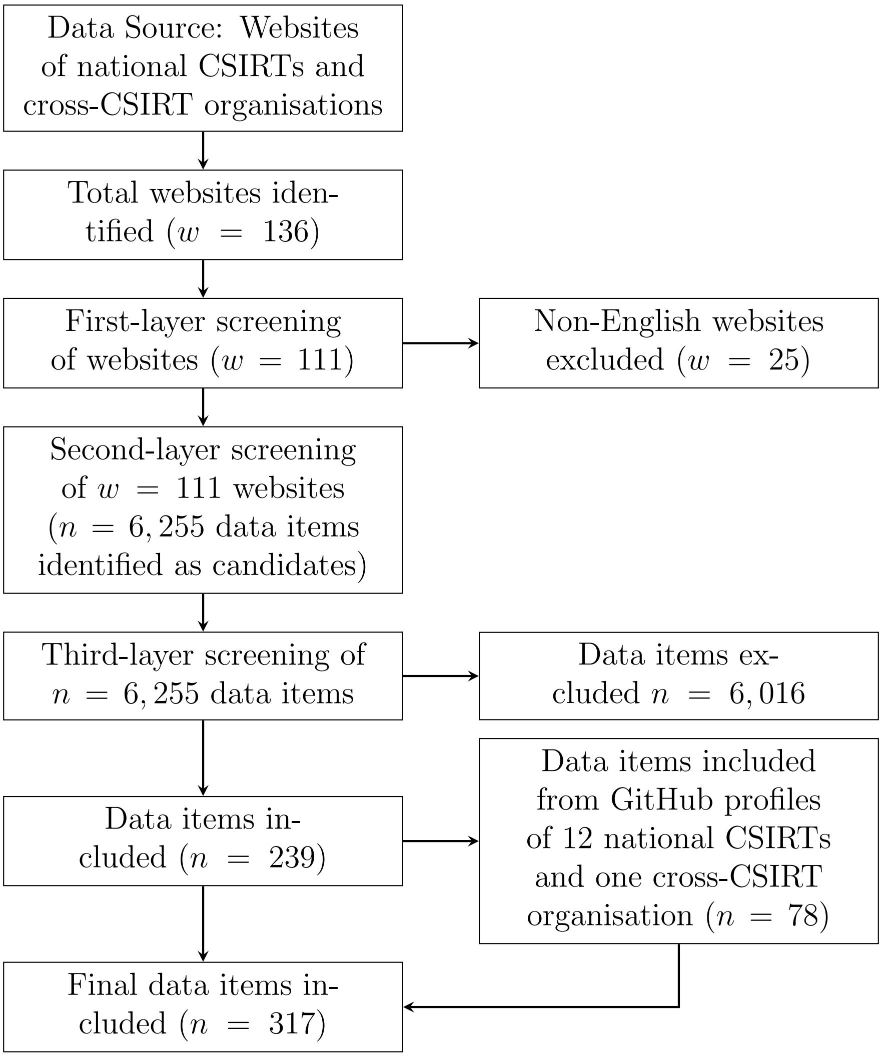

Data Source: Websites of national CSIRTs and cross-CSIRT organisations
Total websites identified ($w = 136$)
First-layer screening of websites ($w = 111$)
Non-English websites excluded ($w = 25$)
Second-layer screening of $w = 111$ websites ($n = 6,255$ data items identified as candidates)
Third-layer screening of $n = 6,255$ data items
Data items excluded $n = 6,016$
Data items included from GitHub profiles of 12 national CSIRTs and one cross-CSIRT organisation ($n = 78$)
Data items included ($n = 239$)
Final data items included ($n = 317$)

Figure 4: Stage 1a: Searching Websites of National CSIRTs and Cross-CSIRT Organisations

### Data Sources

The SLR focused on the publicly searchable portions of the official websites of national CSIRTs and cross-CSIRT organisations. The search identified 125 national CSIRTs from the lists of two key organisations: FIRST[191] and ITU[114]. Non-national CSIRT websites were excluded from the study as they are less public, and their practices are likely more diverse and sector-dependent. The SLR also included 11 cross-CSIRT organisations' websites based on the personal knowledge and experience of the researcher as a national CSIRT staff member. This includes cross-CSIRT organisations identified based on checking the websites of the included national CSIRTs. Hence, the data sources in Stage 1 consist of 125 websites of national CSIRTs and 11 cross-CSIRT organisations, totalling up to 136 websites.

After the first step of screening all 136 websites, 25 non-English websites were excluded, and websites of 100 national CSIRTs and 11 cross-CSIRT organisations for the SLR were included, totalling up to the final 111 websites. The list of national CSIRTs' and cross-CSIRT websites included in the SLR is available at `https://cyber.kent.ac.uk/research/CSIRTs/SLR/List-Websites.html`. The SLR also paid attention to the 100 national CSIRTs and 11 cross-CSIRT organisations and searched if they have an official profile on GitHub (`https://github.com/`). This led to the discovery of 13 GitHub profiles of 12 national CSIRTs and one cross-CSIRT organisation. A Google search was not applied to such GitHub profiles since searching the profiles directly at GitHub was more accurate and more manageable. All public projects of those profiles were checked to identify relevant public data and free tools developed by national CSIRTs for the national CSIRTs community.

A total of 6,255 data items $n$ were identified from the 111 websites (w) of national CSIRTs and cross-CSIRTs organisations. After applying the exclusion and inclusion criteria, 6,016 data items were excluded, and 239 were included. Another 78 data items were identified and included from the GitHub profiles of 12 national CSIRT and one cross-CSIRT organisation; after applying the EC and IC, 317 data items were included for the SLR.

### Keyword Selection

The following keywords were used via Google Search for relevant data items (web pages and online documents) on the 111 national CSIRTs and cross-CSIRT websites. During the search, "[domain name]" is replaced with the domain name of each website, removing "www" when relevant) as in the below search query:

```
site:[domain name] "open data" OR "public data" OR "free data" OR
tool* OR system* OR software OR service*
```

For some websites, the domain name is too broad that the above keyword returned too many pages, so we added the corresponding acronym of CSIRT (e.g., "CERT" for national CSIRTs whose names use "CERT") as a new keyword to limit the returned results.

For each search, all pages returned by Google were manually checked until the following standard message:

> "To show you the most relevant results, we have omitted some entries very similar to the [X numbers] already displayed."

For each returned web page, the partial text returned by Google was read to judge if the web page or online document is likely relevant for the SLR. If so, the web page or the online document is opened for reading in full to check if there is any mention of the use of public data or free tools in the operational practices of one or more CSIRTs.

Links on web pages that are not hosted by the corresponding CSIRT but include some relevant information on public data and free tools were considered candidate data items. Going this additional mile is particularly important for web pages that show a list of public datasets and free tools, e.g., open data at the CIRCL (the Computer Incident Response Team Center Luxembourg, `https://www.circl.lu/opendata/`). Such links include GitHub profiles of national CSIRTs and cross-CSIRT organisations.

### Exclusion and Inclusion Criteria

A rigid set of exclusion and inclusion criteria was not defined in Stage 1. This is because some web pages and online documents only have a small section referring to the use of public data and free tools.

The following exclusion criteria (EC) were used in Stage 1:

1. Websites written in non-English were excluded since English is the primary language the researcher could understand.

2. Web pages that do not discuss or mention the use of public data or free tools by national CSIRTs were excluded.

3. Web pages that discuss general guidelines, recommendations and tools for the general public, Small and Medium Enterprises (SMEs) and general security tools such as penetration-testing and security-testing were excluded.

4. Web pages with too many secondary links, for instance, index pages, were excluded. This is because relevant pages on such pages were usually returned separately by Google search.

5. Web pages that mention common public databases such as CVE (Common Vulnerabilities and Exposures, `https://cve.mitre.org/`), CWE (Common Weakness Enumeration, `https://cwe.mitre.org/`) and NVD (National Vulnerability Database, `https://nvd.nist.gov/`) were excluded. This is because such data sources frequently appear on too many web pages.

After applying the EC, the following inclusion criteria (IC) were applied to select relevant data items for the SLR:

1. Data items that discuss, advocate and mention the use of free tools and public data in national CSIRTs.

2. Data items about tools developed by national CSIRTs, their staff members, and those whose descriptions show relevancy to security incidents and incident responses (e.g., recommended by a national CSIRT or a cross-CSIRT organisation for such purposes).

### Stage 1b – Identifying Relevant Research Papers

Figure 5 depicts the procedure adopted for this stage. A systematic literature search was conducted between 17 and 20 November 2022 on the Scopus Database, one of the most widely used scientific databases of research papers – by applying a keyword search into metadata (titles, abstracts and keywords).

After applying the EC and IC, 22 relevant research papers were included for the SLR. These papers are: [167; 214; 88; 96; 93; 115; 121; 125; 136; 140; 146; 151; 154; 157; 170; 171; 193; 212; 216; 231; 259; 281]. The snowballing technique [280] was applied to search and review the bibliographies of the 22 research papers, adopting the same EC and IC to identify more relevant research papers. Using the snowballing technique, one more research paper was identified [120] and a technical report [32]. In total, 23 research papers and one technical report were included, with a final total of 24 research papers in Stage 1b.

### Data Source

For the data source, the Scopus database, one of the most used and popular scientific indexing databases for searching the research literature, was selected for the study [286]. The Web of Science (WoS) database was also considered, but a search with the Scopus database returned more results than the Web of Science (WoS) database. This is unsurprising when referring to comparative studies of different scientific databases [2; 208; 268]. Furthermore, searching into specific scientific publishers' databases refrained because the Scopus database already covers the majority of key scientific publishers. Searching with

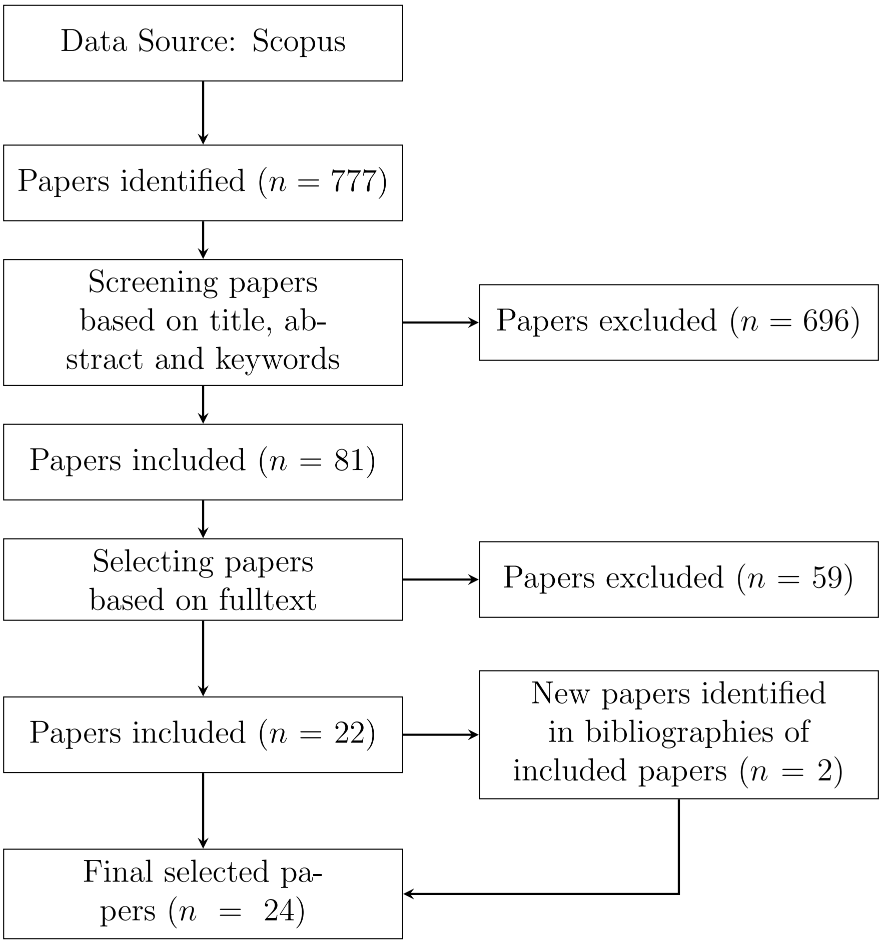

Data Source: Scopus
Papers identified ($n$ = 777)
Screening papers based on title, abstract and keywords
Papers excluded ($n$ = 696)
Papers included ($n$ = 81)
Selecting papers based on fulltext
Papers excluded ($n$ = 59)
Papers included ($n$ = 22)
New papers identified in bibliographies of included papers ($n$ = 2)
Final selected papers ($n$ = 24)

Figure 5: Stage 1b: Searching Research Papers in the Scopus Database

Google Scholar was not considered because it lacked the support of advanced search syntax, such as nested searches and digging into metadata. As an automated search engine, Google Scholar provides less accurate search results and covers too many data items, including papers that are not peer-reviewed and published at less credible venues (e.g., predatory journals and conferences). Google Scholar also returns a limited number of more relevant papers, making it less ideal even if complete coverage is needed[9] .

### Search Keywords

To determine the best search query, various combinations of different keywords were tried and judging on the returned research papers, the search query was finally decided. The search query covered as many relevant papers as possible while returning a manageable number of candidate papers for subsequent screening. The search terms “security OR attack*” were used with “CSIRT” and “CERT” acronyms to eliminate papers that use identical acronyms unrelated to the study topic. The final search query used is shown below:

[9] See `https://harzing.com/resources/publish-or-perish/manual/reference/dialogs/preferences-google-scholar` for some explanation on this search limitation.

```
((cyber* OR computer) AND (incident* OR emergenc* OR security OR attack*)
   AND ("response team*" OR "readiness team*")) OR ((CSIRT OR CSIRTs OR
       CSIRT OR CSIRTs OR CERT OR CERTs) AND (security OR attack*))
```

### Exclusion and Inclusion Criteria

The following EC were used to screen research papers:

1. Papers published before 2010 were excluded to ensure the papers included in the SLR reflect the most recent and current practices of national CSIRTs.

2. Non-English papers were excluded since English is the primary language the researcher could understand.

3. Papers that do not discuss national CSIRTs' operations were excluded.

4. Papers that do not report original work and opinions (e.g., editorials) were excluded.

5. Papers with unclear abstracts (in terms of the quality of technical writing) were excluded because vague or poorly written abstracts typically imply that the results reported are unreliable or difficult to interpret.

After applying the EC, the following IC was applied to select the most relevant papers for the SLR:

1. Papers that cover operational practices of national CSIRTs concerning free tools and public data.

2. Papers that report empirical studies, literature surveys and systematic literature reviews regarding using public data or free tools in national CSIRTs.

3. Papers that report security incident case studies using free tools and public data in national CSIRTs.

4. Papers that propose new approaches to enhance operational practices and capabilities of national CSIRTs using free tools and public data.

### Stage 2 - Data Analysis by Synthesising Data Items from Stages 1a and 1b

Qualitative and minor quantitative approaches were used to systematically review the full text of the documents extracted from national CSIRT websites, cross-CSIRT websites and research papers from the Scopus Database. The data analysis method adopted in the SLR is content analysis, as explained in Section 1.4.2. The content analysis method was used successfully in previous systematic literature review studies [74; 230; 7], motivating similar adoption for the present SLR.

Content analysis was used to get a general overview of the extracted online documents from national CSIRT websites, cross-CSIRT websites, and research papers, reflected in numbers. The method counted the number and frequency of data items appearing in online documents and research papers, illustrated in tables and graphs—for instance, the number of research papers published by year and author country. Content analysis was also used to analyse online documents, web pages, and research papers to identify significant information, describing, explaining, and categorising it objectively rather than interpreting it to answer the RQ. For example, when explaining free tools used in national CSIRTs, the intention is to identify and describe a list of free tools and categorise them into their classes without interpreting them. This includes determining the number of research papers or websites that mention a particular free tool or public data.

### 3.5.3 Results

This section presents Stage 2 of the SLR. It consists of results synthesised from data items identified in Stage 1a (data items from National CSIRT websites and cross-CSIRT organisations) and Stage 1b (research papers). In Stage 1a, 317 data items were identified from the websites of 100 national CSIRTs and 11 cross-CSIRT organisations. In Stage 1b, 24 research papers were identified from the Scopus scientific database.

First, the result presents an overview of the findings synthesised from Stages 1a and 1b reflected in numbers. Second, the result is presented by synthesising findings across the research papers and websites of national CSIRTs and cross-CSIRT organisations to generate overarching explanations that answer the RQ.

#### General Information – from Stage 1b

**Research Papers Published by Year.** The publication year is not always known for the identified web pages and online documents from Stage 1a, so general information regarding their statistics is not shown. Nevertheless, a manual inspection revealed that over three-quarters of web pages and online documents were published within the last five years. A similar observation was found with the research papers published over the previous five years, as shown in Figure 6.

**Research Papers Published by Country.** The majority of the published papers are from the USA (5), followed by Japan (3), Norway (2), Germany (2), Austria (2) and one publication from each of the following countries: Bangladesh, Greece, Indonesia, Finland, Luxembourg, Poland, Serbia, South Africa, Taiwan and the United Kingdom. A summary of the number of research papers published by country is shown in Figure 7.

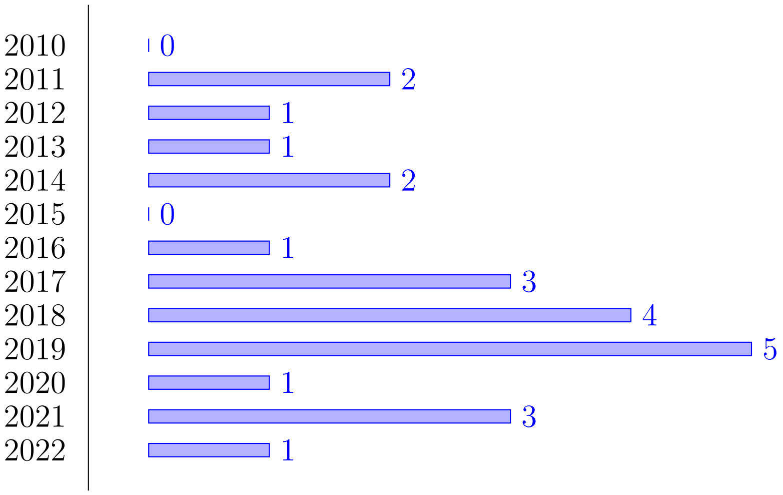


Figure 6: Number of Research Papers Published by Year

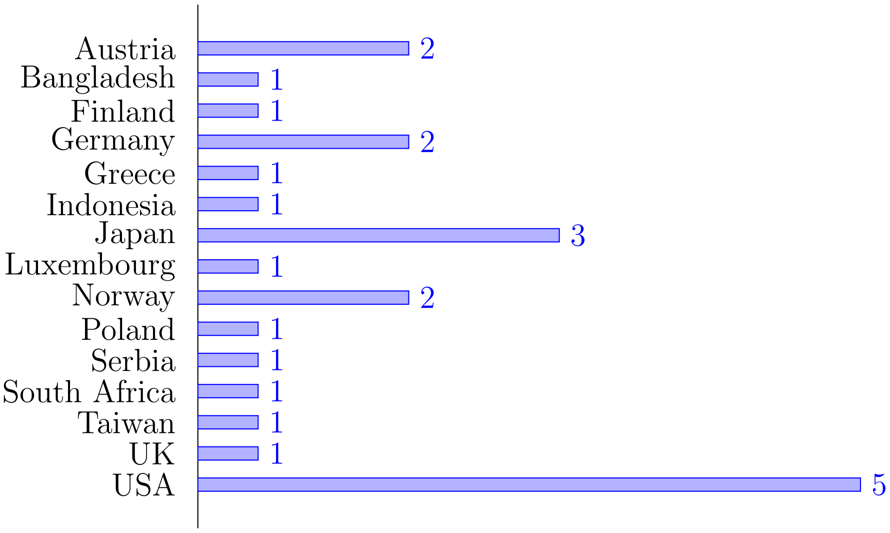


Figure 7: Number of Research Papers Published by Country

**Types of CSIRTs Mentioned in Research Papers.** The national CSIRTs and cross-CSIRT websites covered publications related to national CSIRTs only, whilst the 24 research publications also covered the different types of CSIRTs, as shown in Table 4 considering the lack of research publications on national CSIRTs. Hence, publications that mention other types of CSIRTs' operational practices were included to get a general overview of CSIRTs' operational practices from the research literature.

**Summary of Research Papers Covered in the SLR.** The research papers identified in the SLR covered the operational practices of national CSIRTs and other types of CSIRTs due to the lack of coverage of national CSIRTs in the research publication. The details of the 24 research papers covered in the SLR are summarised in Table 5.

Table 4: Types of CSIRTs Mentioned in Research Papers

| **Type of CSIRTs** | **Research Papers** |
|---|---|
| Regional CSIRTs | [167] |
| National CSIRTs | [146; 170; 32; 167; 96; 88; 214] |
| Public Sector CSIRTs | [214] |
| Organisational CSIRTs | [136; 154; 193; 281; 231; 115; 140; 151] |
| Research Organisation CSIRTs | [157; 125] |
| Academic CSIRTs | [121; 212] |

Table 5: Summary of the 24 Research Papers Covered in the SLR

| **Cite** | **Research Topic** | **Research Method** |
|---|---|---|
| [167] | Personal observation at Malaysia national CSIRT. General information about national CSIRTs, precisely the Malaysia national CSIRT, found a systematic use (e.g. evaluation) of data and tools is lacking. | Case Study |
| [214] | The operational practices of public sector CSIRTs in Germany, and found the analysis of incident data, transparent reporting and tool structures for information exchange needs improvement. | Empirical Study |
| [96] | A proposal to address the lack of incident detection tools, incident response tools and procedures governing the operations of national CSIRT in Bangladesh. | Case Study |
| [88] | Examination of gaps in the national CSIRT of Indonesia by following the Carnegie Mellon University framework and recommended improvement to fill the gaps by referring to the FIRST framework. | Case Study |
| [146] | Examination of citizens' and organisations' roles within the formal cyber security networking to help identify and report incidents in Finland. | Nodal cluster analysis |
| [115] | Investigation of cyber incident response readiness and capacity in the Norwegian petroleum industry for handling critical security incidents. | Empirical study |

Table 5 – *Continued from previous page*

| **Cite** | **Research Topic** | **Research Method** |
|---|---|---|
| [140] | Investigation of general machine learning applications in cyber security and how machine learning can be adopted to improve operations in CSIRTs. | Literature survey |
| [154] | Description of the European Air Traffic Management CSIRT's functions, services and partnerships with other CSIRTs on threat information sharing to detect, respond, mitigate and prevent cyber threats. | Case study |
| [170] | Development of an automated tool, the AIL, to detect information leaks in Luxembourg by crawling paste websites on the Internet to facilitate incident responses. | Tool development |
| [281] | Investigation of the practices of automatic information exchange among CSIRTs regarding the tools, approaches, standards, challenges and technologies. | Systematic literature review |
| [121] | Investigation and incident response to an information leak at a University in Japan, conducted by the university's CSIRT. | Case study |
| [136] | Integration of cyber threat intelligence to a corporate-based security framework, i.e. NTT CSIRT of Japan, to defend networks and systems against sophisticated cyber attacks. | Case study |
| [193] | A forensic-based incident response in NEC CSIRT of Japan using an in-house developed tool to preserve data and evidence related to incidents. | Case study |
| [93] | A GDPR compliance checking model and tool for supporting the information exchange between CSIRTs, to be integrated into the CSIRTs' information sharing process. | Model and tool development |

Table 5 – *Continued from previous page*

| Cite | Research Topic | Research Method |
|---|---|---|
| [212] | Automating information exchange at an Academic CSIRT for real-time identification of new cyber threats and cyber-attacks. | Business Intelligence applications and Online Analytical Processing with Ralph Kimball methodology. |
| [259] | Development of early warning alerts of upcoming malicious activities for CSIRTs to improve overall incident response. | A Data Warehouse and Business Intelligence applications, with Ralph Kimball methodology. |
| [171] | Designing an improvised model for establishing CSIRTs, which addresses the deficiency of a standard model for developing CSIRTs | Design Science Research |
| [231] | A critical review of state-of-art and considerations when building effective security information sharing platforms for CSIRTs. | Literature survey |
| [216] | Establishment of CSIRTs, its objectives to deal with cyber attacks, evaluating CSIRTs' performance and information sharing services in CSIRTs. | Case study |
| [120] | Describes and discusses a list of options on information sharing for CSIRTs and clarifies misconceptions about information sharing. | General discussions |

Table 5 – *Continued from previous page*

| Cite | Research Topic | Research Method |
|---|---|---|
| [32] | Investigation of existing communication solutions and practices among European CSIRTs, and the gaps that limit cyber threat information exchange in Europe. | Literature survey and interviews |
| [151] | Introducing Nagios, a tool for monitoring known vulnerabilities and Pakiti, a tool for systematically patching vulnerable systems in the EGI CSIRT network. | Case study |
| [125] | Introducing several in-house projects and systems for automatically detecting cyber attacks at CSIRTs in Poland. | Case study |
| [157] | Orchestration of incident reporting capabilities, automated incident responses, and process-oriented intervention for integrated management of security incidents | Case study |

### Significant Information Identified from the SLR

Three significant pieces of information directly answer the RQ of the research, whilst three others are not related to the RQ but are considered for the research as they reflect national CSIRTs. These are listed below:

1. The Use of Free Tools in National CSIRTs' Operations
2. The Use of Public Data in National CSIRTs' Operations
3. How National CSIRTs' Staff Perceive Public Data and Free Tools
4. Active Development and Advocating of Public Data and Free Tools by National CSIRTs and Cross-CSIRT Organisations
5. Information Sharing Among National CSIRTs
6. Challenges and Issues with Tools and Data in National CSIRTs' Operational Practices

**1. The Use of Free Tools in National CSIRTs' Operations.** The CSIRT publications and research papers mentioned free tools, open-source tools, closed-source tools, freeware and free online services used in national CSIRTs and CSIRTs operations. The websites of national CSIRTs, Cross-CSIRT organisations' and their respective GitHub websites mentioned 301 free tools to facilitate incident responses for CSIRTs. The free tools mentioned can be classified into the following types:

1. Log Analysis
2. Artefact Analysis
3. Incident Management
4. Whois
5. Web Information Gathering
6. Network Monitoring
7. Sandbox and Reverse Engineering
8. Penetration Testing
9. Parser
10. Scanner
11. Disk Image Creation
12. Evidence Collector
13. Visualisation
14. Cyber Threat Intelligence
15. Honeypot
16. Memory Analysis tool
17. Search Engine
18. File Recovery
19. Linux Distributions
20. Communication tools
21. Network Utility
22. Intrusion Detection System

From the SLR, it is understood that the developers of the free tools consist of software companies, researchers affiliated with research organisations and CSIRTs, independent researchers, independent developers, CSIRT staff, and joint projects or collaborations between CSIRTs and researchers. The free tools are generally recommended for national CSIRTs', CSIRTs and broader security operations to facilitate incident response.

The research papers mentioned 64 free tools. These are incident response and management tools (i.e. investigative tools, ticketing systems, reporting systems), intrusion detection systems, information sharing and exchange platforms and OSINT tools. The researchers developed some of these tools, while others are third-party tools to facilitate research reported in the papers.

Tools identified from the research papers can be classified into different types, shown in Table 6.

Table 6: Tools Identified from Research Papers

| **Class of Tools** | **Example Tools with Citation** |
|---|---|
| SIEM | iView [231; 96], OSSIM [157], SIEM [96] |
| Vulnerability scanners | nmap [157], deep exploit [140], OpenVAS [121] |
| Information sharing | MISP [170; 281; 231; 93; 154; 214], CIF [231], TAXII [231; 93] |
| Incident management | RTIR [170], TheHive [170; 32], OTRS [214] |
| Incident reporting channels | Email, telephone [214] |
| Network monitoring | Autoreport [146], Nyx [157], Nagios [151] |
| Artefact analysis | Cuckoo sandbox [140], Nfdump [157], LogHound [231], YARA [120] |
| Communication channels | Email [32; 214], Skype [32], IRC [115], Chat [214], Wiki page [214] |
| Intrusion Detection System | Snort [212; 157], IDS [214] |
| Web information gathering | AIL framework [170] |
| Social media monitoring | ScatterBlogs [214] |
| Cyber Threat Intelligence | MISP [170; 281; 231; 93; 154; 214] |

A comprehensive list of 299 free tools, identified in this SLR, is compiled and made available as an open resource for practitioners, available at `https://cyber.kent.ac.uk/research/CSIRTs/SLR/List_Open-source_Free_Tools.xlsx`.

Several researchers highlighted the significance of tools and data in facilitating incident responses in CSIRTs. For instance, Tondel et al. (2014) conducted a systematic literature review of incident response practices of various organisations [250]. The authors highlighted the importance of tools support and automation to facilitate incident response. The systematic literature review findings are similar to a case study by Kijewski and Kozakiewicz (2011), who examined the operations of CERT Polska (a CSIRT under the Research and Academic Institution) [125]. The authors found tools and data are of utmost importance in CERT Polska, such as threat detection tools, honeynets, and classified and closed-source data, to detect and respond to cyber threats more efficiently and timely.

Mana and Friligkos (2019) experiences and personal observation informs the active use of automated tools such as Malware Information Sharing Platform (MISP) in the operational practices at Eurocontrol Air Traffic Management CSIRT (EUROCONTROL/EATM-CERT) [154]. The authors also revealed that the automated tools allow the active sharing of threat data and other threat information among internal CSIRTs within the EATM-CERT and with several other national CSIRTs in Europe efficiently and timely.

Metzger et al. (2011) reported a case study of the operational practices at the Leibniz Supercomputing Centre (LRZ-CSIRT), an Academic CSIRT in Germany [157]. The authors provided a comprehensive understanding of the operational practices of incident response at LRZ-CSIR, which uses various tools to support incident response. This consists of orchestrating reporting capabilities, automatic analysis and response, and process-oriented intervention for integrated management of security incidents.

Hossain et al. (2021) highlighted the Security Information & Events Manager (SIEM) tool as essential for monitoring and incident response and management purposes in mitigating threats for national CSIRTs [96]. Besides, the SIEM tool displays the data collected by network and security systems for visualisation and helps to prioritise the workload in national CSIRTs.

Riebe et al. (2021) highlighted the use of the OTRS[10] tool for ticketing systems, email and telephone for incident reporting and management in public sector CSIRTs in Germany [214]. Excel sheets were used in the past to manage incidents in public sector CSIRTs in Germany. IDS are also used in Germany's public sector CSIRTs to detect and identify incidents. MISP is used for cyber threat intelligence sharing, while ScatterBlogs[11] is used for monitoring and analysing social media information.

A noteworthy point highlighted by all the above authors is the importance of tool support in the operations of CSIRTs to facilitate incident responses. Hence, there is a need for constant progress in security-related research, with attention explicitly to tools that could help increase efficiency in CSIRTs' operations. Kijewski et al. (2011) stated that "as our lives are increasingly dependent on the Internet and information systems, network threats have a massive potential for impact and become more widespread, necessitating constant progress in security-related research" [125].

**2. The Use of Public Data in National CSIRTs' Operations.** National CSIRTs rely on data from various sources and tools for daily incident responses. National CSIRTs fuse public data with closed-source to help enrich data about threats and facilitate effective incident responses. Commonly used closed-source data include incident reports, cyber threat intelligence data from trusted partners, and closed-source security feeds from trusted sources provided to national CSIRTs for botnets, malicious and phishing website take-downs. Public data mainly refers to data publicly available on the Internet, free for

[10] See `https://www.otrs.com/`

[11] See `https://www.scatterblogs.com/`

download without restrictions, and often obtained via Open Source Intelligence (OSINT) tools.

More than three-quarters (263) of the web pages and online documents identified in Stage 1a of the SLR mentioned free tools compared to public data. To further understand if these free tools utilise public data, the official websites of all the free tools were looked through. It was found that some tools use a comprehensive list of public data, e.g., IntelMQ, which reads from many public data sources (`https://intelmq.readthedocs.io/en/latest/user/feeds.html`), whilst others do not. While national CSIRTs use various public data, they also create and maintain some. These public data can be classified as below:

1. **Data about known vulnerabilities and exposure**: CIRCL CVE daily JSON dump (`https://cve.circl.lu/static/circl-cve-search-expanded.json.gz`)

2. **Security feeds and threat intelligence data**: CIRCL Images Phishing Dataset (`https://www.circl.lu/opendata/circl-phishing-dataset-01/`), CIRCL Images AIL Dataset (`https://www.circl.lu/opendata/circl-ail-dataset-01/`), IOC-DB (`https://labs.inquest.net/iocdb`) and MITRE ATT&CK (`https://attack.mitre.org/versions/v8/software/`)

3. **Data about domain registration information**: WHOIS Data (`http://https://lookup.icann.org/`), RIPE database (`https://www.ripe.net/manage-ips-and-asns/db`), CERT.at Taxonomy of Domain Names Labelling (`https://github.com/certat/awesome-taxonomyzoo-list`)

4. **General-purpose data**: Government of Serbia open data (`https://www.ite.gov.rs/tekst/en/30/open-data-portal.php`), Latvia institutions open data (`https://data.gov.lv/eng/about`)

On the other hand, few research papers from the literature have comprehensive coverage concerning public data in national CSIRTs and CSIRTs in general. Several (8) research papers did indicate that public data is used in CSIRTs' operations, but information about what type of public data is used and how it is used and perceived is largely missing – indicating a research gap.

The public data identified from research papers can be conceptually categorised as follows:

1. **Data about known vulnerabilities and exposure**: Common Vulnerabilities and Exposures (CVE, `https://cve.mitre.org/`) [170; 151], National Vulnerability Database (NVD, `https://nvd.nist.gov/`) [121]

2. **Security Feeds and Cyber Threat Intelligence datasets**: Team Cymru Hash Registry (`https://team-cymru.com/community-services/mhr/`) [32], data from Malware Capture Facility Project (`https://mcfp.weebly.com/`) [212], Spamhaus blocklists (`https://www.spamhaus.org/`) [32], Zone-H.org (a crowdsourcing-based dataset of defacement attacks) [259], malwaredomains.com [125], Shadowserver(`https://www.shadowserver.org/`) [214; 167]

3. **Data on general-purpose at public websites**: public media news [146], posts on public web forums [231], data on online social networks (OSNs) such as Twitter [121], TweetDeck [214], Manufacturer websites [214]

4. **Data at public websites with a strong ICT flavour, not for cyber security per se only**: data shared on websites such as Pastebin.com [170; 136], gist (`https://gist.github.com/`) [170] and codepad (`http://codepad.org/`) [170], data on Alexa.com [125], Security Advisories [214]

5. **Publicly accessible data on darknet** [136; 170]

6. **Open-source intelligence (OSINT)** [214; 167]

7. **Cyber range data** [96]

The two lists above on public data are neither complete nor representative of national CSIRTs due to the small number of research papers the SLR covered and the ad hoc mentions of public data in the national CSIRT and cross-CSIRT websites. It also appears that national CSIRTs may not have utilised some public data sources online due to a lack of systematic information about such data and how to search them easily.

Security organisations also share "restricted data" with national CSIRTs for national-level take-downs of botnets and phishing websites, considering national CSIRTs as the national points of contact. Such data is not available to the public. This is exemplified by the Shadowserver Foundation[12] , which shares data on botnet infections to national CSIRTs and other trusted entities for national-level eradication of botnets. Team-Cymru[13] also works with national and regional CSIRTs worldwide by sharing similar data, such as threat intelligence, so national CSIRTs are well informed of the threats in their constituencies.

**3. Active Development and Advocating of Public Data and Free Tools by National CSIRTs and Cross-CSIRT Organisations.** The SLR study found several national CSIRTs are actively developing and promoting public data and free tools. This is reflected in the data items identified in the SLR, summarised in Table 7 and Table 8.

[12] See `https://www.shadowserver.org/`

[13] See `https://www.team-cymru.com/`

CIRCL is one of them, a national CSIRT from a small country in Europe that significantly impacts the national CSIRT community.

Table 7: Tools Developed by National CSIRTs

| **National CSIRTs** | **Tools Developed** |
|---|---|
| CIRCL (Luxembourg) | MISP, AIL Framework, Circlean, BGPranking, URL-abuse,Potiron, Carl-Hauser, D4 Attack Map, DEFT, Douglas-Quaid |
| CERT.at (Austria) | IntelMQ, CERTspotter-processing, IP2nat, FollowTcpStream |
| JPCERT/CC (Japan) | MalConfScan, aa tool, Sysmon Search, Logon Tracer, Emocheck, MalConfScan-with-Cuckoo, DetectLM |
| INCIBE-CERT (Spain) | BotChecker Script, Terminology extractor, Merovingio, Onion Indexer |
| US-CERT (USA) | CHIRP IOC Detection Tool, Sparrow, Aviary |

Many national CSIRTs also actively recommend various free third-party tools to the national CSIRT community that could help facilitate incident responses. Apart from tools, national CSIRTs also develop and advocate public datasets for the CSIRT community and the general public, as shown in Table 8.

Table 8: Public Dataset Developed by National CSIRTs

| **National CSIRTs** | **Public Dataset Developed** |
|---|---|
| CIRCL (Luxembourg) | CIRCL Images Phishing Dataset, CIRCL Images AIL Dataset on Information leak, Allaple malware infection raw data |
| CERT.at (Austria) | Taxonomy of Domain Names Labelling, Internet-inventory of metadata on IPs and networks and ASNs on the net |
| JPCERT/CC (Japan) | Cyber green database of cyber security risks and vulnerabilities |
| INCIBE-CERT (Spain) | ICARO database of IOCs |
| SWITCH-CERT (Switzerland) | Connectcome Knowledge of scientific data across multi-disciplines |

Some research papers advocate tools such as a Security Information & Events Manager (SIEM) tool, essential for national CSIRTs to facilitate monitoring and incident response and management [96]. It was also stated that it is critical for national CSIRTs to identify the infrastructure and tools needed to support the operations by prioritising open technology (open-source tools), which is free, independent and cost-saving [88].

**4. How National CSIRTs' Staff Perceive Public Data and Free Tools.** None of the data items identified from national CSIRT and cross-CSIRT publications explicitly indicate how national CSIRT staff perceive the usefulness of public data and free tools for operations. However, the fact that several national CSIRTs actively develop and advocate public data and free tools implies a positive attitude and perception of public data and free tools for the operations of national CSIRTs. The fact that many national CSIRTs (16) use free tools, inclusive open-source tools, gives the impression that free tools are perceived positively by national CSIRTs' staff. This finding is consistent with another study, which found that support for open-source tools is solid in the national CSIRT community [59]. This can be seen from the most advanced and popular tools used in the CSIRT community, which are open-source tools developed by national CSIRTs [59]

Similarly, none of the research papers has empirical evidence of explicit and direct coverage of how national CSIRT staff perceive public data and free tools. Nevertheless, the usage of free tools and public data in national CSIRTs [170; 157; 154; 121; 193; 167; 214] indirectly implies the staff's sound acceptance of free tools and public data in their operations. Staff satisfaction from other types of CSIRTs (non-national CSIRTs) reflected in surveys and interviews from previous studies [115; 259; 151; 32] shows CSIRT staff's positiveness towards public data and free tools, in general.

**5. Information Sharing Among National CSIRTs.** According to national CSIRTs' and cross-CSIRT websites, national CSIRTs actively collaborate and share information, threat data and tools (free and open-source tools) within the national CSIRTs' community. This is through initiatives and collaborative projects, primarily by national CSIRTs and cross-CSIRT organisations advocating information sharing. For instance, the MeliCERTes Cyber Security Platform initiative [69] is a modular platform co-developed by several national CSIRTs in Europe. MeliCERTes allows national CSIRTs to share information and collaborate within a verified trust circle of national CSIRTs. The EU-supported CyberExchange project [62] is another initiative to exchange information, knowledge and skills collaboratively among European national CSIRTs. Furthermore, national CSIRTs worldwide also collectively exchange information on indicators of compromise (IOCs) through MISP. MISP is an IOC information-sharing platform co-developed by several national CSIRTs from Europe.

Cyber-attacks now cross boundaries between countries and need international collaboration to respond and mitigate, necessitating comprehensive and effective global and regional collaboration [96]. Global cooperation among international security players, including national CSIRTs, is vital for sharing cyber security information [93] and operational information about threats [125] for quick detection and responses to incidents. Information sharing among CSIRTs, through private-public collaboration and international cooperation in a more structured manner, is essential for national CSIRTs [231; 136]. Such collaborations can be achieved through closed “mailing lists” or “encrypted email” communication, aligned with standard formats for information sharing [93] and automated information exchange such as by MISP [170; 281; 231; 154], CyBOX, CYBEX, STIX and TAXII [281; 32; 171] between CSIRTs, national CSIRTs and security communities.

Different protocols and standards are in place to regulate information sharing, for instance, the Traffic Light Protocol (TLP) [190; 257]. The TLP determines the confidentiality of information during communication by classifying documents or information as red, amber, green, or white. The TLP is necessary when sharing information with CSIRTs about intelligence, indicators of compromise, incident reports, vulnerability information and malware detection [154; 281; 32; 115]. Despite the current information-sharing practices among national CSIRTs, the effectiveness is questionable [281]. Notably, technical issues were identified as the most common barrier to effective information exchange among CSIRTs [32].

**6. Challenges and Issues with Tools and Data in National CSIRTs’ Operational Practices.** Issues and challenges concerning national CSIRT operational practices are less known on national CSIRTs’ websites and cross-CSIRT websites. Nonetheless, a few research papers mentioned issues and challenges in the operational practices of national CSIRTs and CSIRTs. Issues concerning the quality and usability of tools and data were pointed out by Werlinger et al. (2010) [274] – indicates an operational gap. The authors discovered the issue from an exploratory study of incident response practices in the academic, governmental, and private sectors of CSIRTs. The authors highlighted that practitioners often need to develop in-house tools for diagnostics and detecting incidents, as an incident response is highly collaborative work.

This means that national CSIRTs currently do not have access to quality and usable free tools and data for adoption in operations. Furthermore, the detection of incidents is becoming complicated, requiring expertise on the part of the staff and effective support from qualified, usable and reliable tools. The point highlighted by Werlinger et al. (2010) is crucial and requires attention from researchers and stakeholders.

Similarly, Bourgue (2013) highlighted the lack of quality data and software (tools) to support effective incident responses in CSIRTs, calling for more research that focuses on tools and data, precisely on the quality and usability aspects [32]. While Werlinger et al. (2010) revealed the issue with general incident response tools, Metzger et al. (2011) specifically highlighted the issue of forensic tools used in the CSIRT. Regardless of the tools’ type, a crucial point the authors are trying to convey is the issues with tools and data used for incident responses in national CSIRTs and CSIRTs that need to be addressed with prospective research.

Several research papers [140; 121; 125; 157] also highlighted the concerns in CSIRTs on automation and advanced tools to facilitate effective detection and analysis of security incidents. This is because manual handling of incidents is a challenge to CSIRTs as it is prone to human errors causing loss of incident evidence, time-consuming and requires more resources [193]. This can be seen when dealing with "Information overload" on vulnerabilities and attacks, as more time and resources are needed to extract relevant information for incident responses [115]. Other challenges highlighted in the research papers include insufficient security experts and qualified personnel in CSIRTs [140; 121; 193; 216], lack of logging security events and network monitoring [115] and lack of systematic preventive measures against cyber attacks [157]. Some challenges related to the development of tools for incident responses have been addressed [125; 170; 212] whilst many other challenges in CSIRTs and national CSIRTs remain unsolved.

## 3.6 Research Gaps

Notably, the SLR identified several research gaps, highlighted and summarised in the following points; 1) lack of research and public reports on the understanding and adoption of public data and free tools in national CSIRTs' operations; 2) systematic evidence on the benefit of using free tools and public data to facilitate incident response in national CSIRTs is lacking; 3) lack of studies on the development of systematic procedures concerning the use of public data and free tools in facilitating efficient incident response; 4) what most national CSIRTs are doing with public data and free tools in the operations is less known from the research literature compared to national CSIRTs and cross-CSIRTs publications; 5) the diversity of free tools and public data seems insufficient, with a lack of understanding of what is needed and what is still missing in national CSIRTs; 6) none of the national CSIRT publications and research papers has empirical evidence of direct coverage on how national CSIRT staff perceive public data and free tools are useful in the operations.

The key findings and research gaps identified from the SLR suggest the need for more empirical research regarding the use of free tools and public data in the operational practices of national CSIRTs. Challenges raised by some research papers must not be taken lightly; future work must address these challenges, particularly those involving tools and data quality, including usability. Though there are studies on cyber security threats and their countermeasure, very limited studies are available concerning national CSIRTs [96]. Such research is crucial for formulating systematic procedures for national CSIRTs and enhancing the current operational practices [88]. Therefore, more empirical research is recommended on national CSIRTs, which have been studied less in academic research. This includes the development of standardised and systematic procedures for critical areas of operational practices at national CSIRTs (e.g., for tool and data evaluation), subsequently addressing some of the issues raised in this study. This suggestion is reinforced by previous studies recommending more research concerning tools, tool development and evaluation in CSIRTs' operations [250]. Furthermore, research about tools [115] and how technology can best provide support for security incident responses [274] can help to address the lack of advanced tools for analysis work in CSIRTs [261].

## 3.7 Summary

This chapter presented a literature review of the research topic, followed by an SLR of the operational practices of national CSIRTs regarding the use of free tools and public data. This chapter provides a holistic literature review of the research topic. The SLR was performed and presented to give an in-depth understanding of the operational practices of national CSIRTs. The results of the SLR were described in detail. Several research gaps identified from the SLR are outlined with suggestions to advance prospective future work to address these gaps.

It should be noted that the SLR only provided a limited understanding of staff's perceptions, so a more in-depth understanding was intended to be achieved using empirical studies. Furthermore, a lack of open discussions is observed on how national CSIRTs use and perceive public data and free tools to facilitate cyber incident responses. Hence, empirical studies with multiple national CSIRTs worldwide are necessary to gain insights and better understand real-world operations. This is precisely about the use of free tools and public data, how staff perceive free tools and public data for the operations, and some solutions that can be proposed for national CSIRTs. Therefore, an empirical study to gain real-world insights into national CSIRTs' operations is undertaken and presented in the next chapter to follow up on the current chapter.

# Chapter 4

# Empirical Study to Gain Insights into the Real-world Operational Practices of National CSIRTs Regarding the Use of Free Tools and Public Data

This chapter is a follow-up to the Systematic Literature Review reported in the Literature Review Chapter 3. This chapter presents an empirical study of real-world operational practices of national CSIRTs concerning the use of free tools and public data and how such tools and data are perceived across national CSIRTs – a topic less understood from the literature. The study was conducted to answer RQ2 of this research defined in Section 1.3. This chapter also introduced Open Source Intelligence (OSINT) in the later stage of this study, which was not covered in the previous chapter. The study presented in this chapter was conducted in two phases, with a single national CSIRT in the first phase, followed by multiple national CSIRTs in the second phase. The chapter is structured as follows. Section 4.1 introduces the empirical study and the original contributions. Section 4.2 explains the methodology adopted for the study, primarily the data collection and analysis. Section 4.3 provides the results and findings of the study, and Section 4.4 summarises the chapter.

## 4.1 Introduction

National CSIRTs typically have access to data and investigative tools from various sources to support their operations. This includes closed-source data of self-reported incident data or evidence from victims (including organisations and citizens), law enforcement agencies (LEAs), and trusted partners such as collaborating CSIRTs. However, these closed-source data and tools are often insufficient for national CSIRT staff to conduct investigations effectively. For instance, the researcher, a Malaysia Computer Emergency Response Team (MyCERT) staff member, observed that the availability of closed-sourced data and commercial investigative tools in national CSIRTs is limited [167]. This is due to the insufficient closed-source data that organisations provide when reporting incidents, while commercial tools are too costly. Therefore, national CSIRTs often resort to public data and free tools available on the Internet for additional resources to support incident responses [167].

Public data has long been used to enrich and complement closed-source data in many applications of various fields, e.g., to facilitate crime investigation by Law Enforcement Agencies (LEAs) for timely, reliable and actionable intelligence [51; 273]. Public data is also used for efficient business decision-making to increase product sales [67] and monitor virus spread in medical situations [188; 150]. Digital forensics uses public data to obtain intelligence because public data is "fast, flexible, dynamic, communicable, shareable and partner forming" [210].

OSINT tools, free tools, open-source tools, and free online services have great potential to contribute to incident responses, such as collecting intelligence about potential cyber threats, in-depth information about ongoing incidents, and other meaningful information and data. OSINT tools, open-source tools, free tools and free online services can even be used to inspect threats against critical infrastructures [143], to track cyber criminals' activities at an early stage [199] and to simulate how cyber criminals would conduct cyber attacks – an understanding of which would help mitigate such threats [255]. Furthermore, such tools can improve cyber security posture in organisations [215] to guide searching novel information about a threat [77]. Interestingly, OSINT tools, open-source tools, free tools and online services are also used for background checks during job hire and to verify the authenticity of students' assignments [198].

A systematic literature review conducted as part of this research and reported in Chapter 3 indicated a lack of systematic open discussion concerning how public data and free tools are used and perceived in national CSIRTs' operations. Hence, the present study is undertaken to understand better the operational practices of national CSIRTs regarding the use of OSINT, free tools and public data and how such tools and data are perceived across real-world national CSIRTs. Notably, the study is also interested in understanding another type of tool, the OSINT tool, which was not covered in the systematic literature review. OSINT tools are complicated and use public data; hence, a better understanding of OSINT tools from the empirical study is sought.

The study was conducted in two phases. In the first phase, the study was conducted between 21 July and 14 August 2020 with a single national CSIRT, the Malaysia Computer Emergency Response Team (MyCERT). In the second phase, the study was conducted between 9 October and 27 November 2020 with 12 other national CSIRTs worldwide. The second phase helped to enlarge, compare and validate findings from the first phase with viewpoints and perspectives from a more diverse set of national CSIRTs. This helped to avoid biases observed from a single national CSIRT (MyCERT) for credibility and ensured the findings from the study were more representative of national CSIRTs'. MyCERT was selected in Phase 1 as the researcher is an employee of MyCERT, with direct access to the gatekeeper, the Head of MyCERT.

The results from the study presented in this chapter led to five main findings, summarised as follows:

1. Empirical evidence into how OSINT tools, free tools and public data are used in real-world operational practices of the surveyed and interviewed national CSIRTs to facilitate incident response.

2. Empirical evidence confirming that national CSIRTs staff who participated in the study perceived OSINT tools, free tools and public data as useful in facilitating incident response in the real-world operations of national CSIRTs.

3. The insights into how OSINT tools, free tools and public data are used in national CSIRTs can be adopted by other national CSIRTs and the broader security operations.

4. List of OSINT tools, free tools, open-source tools, and public data identified from this study could be a reference for other national CSIRTs and broader security operations.

5. Gaps identified from the study could guide prospective future research and development in this topic.

The rest of this study is organised as follows. Section 4.2 explains the methodology used in this study – data collection and data analysis, and Section 4.3 presents the results.

## 4.2 Methodology

It should be noted the study was conducted in two phases. In the first phase, the study was conducted between 21 July and 14 August 2020 with a single national CSIRT – the Malaysia Computer Emergency Response Team (MyCERT). In the second phase, the study was conducted between 9 October and 27 November 2020 – with 12 other national CSIRTs worldwide.

The primary approach used in this study is qualitative. A minor quantitative approach was used to get an overview of the data by translating it into numbers and figures. The quantitative findings inform the qualitative, where the key findings of the overall study are derived. This study received a favourable opinion from the University of Kent Ethics Board under reference number (CREAG:621920). Two ethics approvals were granted for this study: on 7 July 2020 for Phase 1 Study and 7 October 2020 for Phase 2 Study, while maintaining the same ethics reference number – (CREAG:621920).

Participants were provided with a Participant Information Sheet (PIS) with details of the study and assurance that their personal identifier information (PII) would be protected. Participants were given Consent Forms to consent to participate in the online survey and semi-structured interviews. The PIS and the Consent Form were attached to the online survey, which was emailed to participants.

All participants willingly consented to participate in the research, record their conversations, and include direct quotations in research publications without disclosing their personal information. For privacy purposes, participants' identities were anonymised, and to provide context for specific quotes, pseudonymous identity documents (IDs) were used. Participants' affiliations were also anonymised if requested by participants. The Consent Form used in this study is available at (`https://cyber.kent.ac.uk/research/CSIRTs/Interview-Survey/Consent-Form.pdf`), and the Personal Information Sheet (PIS) used in this study is available at (`https://cyber.kent.ac.uk/research/CSIRTs/Interview-Survey/PIS.pdf`).

### 4.2.1 Data Collection

The data collection method used in this study is an online survey and online semi-structured interview. The survey aimed to collect participants' demographic information, how they investigate incidents, their experience using free tools and public data, and their perception of free tools and public data. The semi-structured interviews aimed to elicit detailed insights from the ones provided in the survey, from which significant findings of the study are derived. The semi-structured interviews also helped clarify some information given by participants in the survey, with more details and examples, improving the quality of data collected in this study. Previous studies had used a similar approach to understand CSIRTs' operational practices [82; 97; 183], motivating the same method for this study. In this study, an online survey was conducted first, followed by semi-structured interviews. It should be noted that the study was conducted in two phases: Phase 1 with MyCERT and Phase 2 with 12 other national CSIRTs. In each phase, data collection was done in two stages: Stage 1 – an online survey, followed by Stage 2 – online semi-structured interviews.

**Survey.** Surveys are generally conducted over the phone or through paper-based questionnaires. For convenience, this study adopted an online survey for data collection [217].

Researchers have widely used surveys (especially online) in various fields, including cyber security, to collect useful information from recruited human participants [79; 222; 123; 21; 37]. In this study, a web-based online "survey questionnaire" was optimised as the instrument for quick, cost-effective and flexible data collection [21]. Moreover, the study's participants are dispersed worldwide, so an online survey is considered the best option for this study [37]. Considering the security and privacy of data collection, the survey was conducted using Typeform[14] , a subscription-based online survey platform commonly used by researchers at the University of Kent. The online survey for this study was conducted between 20 July 2020 and 16 November 2020.

The survey questionnaire consists of 23 questions in three sections:

1. Section A is about participants' work experience and job roles in the national CSIRT to help contextualise their answers to the questions and what they said in the follow-up interview.

2. Section B is about the type of OSINT tools and data used in the operational practices of national CSIRTs and how useful they are, as perceived by staff.

3. Section C tries to capture information on operational challenges at national CSIRTs concerning the use of OSINT tools and public data.

The online survey questionnaire contains some questions that have multiple choices, but often with an "Others" option and an open-ended text box for participants to fill in further details. Some other questions are entirely open-ended, so participants can fill in what they see fit. The online survey questionnaire is available at (`https://cyber.kent.ac.uk/research/CSIRTs/Interview-Survey/Survey-Questionnaire.pdf`).

**Semi-structured interview.** Semi-structured interviews allow detailed insights to be drawn from participants through interactive discussions [187]. Semi-structured interviews were used for this study as they are more suitable for gaining detailed insights, asking complex questions, and exploring experiences and feelings that cannot be obtained through survey questionnaires [187]. This aligns with the study's aim to collect as many details as possible from the participants to draw out more facts. Interviews provide more flexibility for asking questions and discussing issues. The instrument used in the semi-structured interview is an "Interview Schedule" to guide the interviews. It contains brief guidelines on the interview process and interview questions. The Interview Schedule consists of 25 questions about participants' general information, how OSINT, free tools and public data are used by participants and the challenges they face in their work. A copy of the Interview Schedule used for the study is available at (`https://cyber.kent`

[14] See `https://www.typeform.com/`

.ac.uk/research/CSIRTs/Interview-Survey/Interview-Schedule.pdf). The order of the interview questions was flexible, whereby interviewees were allowed to highlight or introduce any relevant points related to the interview questions. The interviewees are free to speak at any length or detail or raise new issues if they think it is appropriate to the interview questions.

The online semi-structured interviews were conducted between 6 August 2020 and 27 November 2020. Whatsapp Call[15] , Microsoft Teams[16] , Zoom[17] and Google Meetings[18] were used to conduct the online semi-structured interviews as preferred by the interviewees. Six interviews were conducted using WhatsApp calls, one through Google Meetings, one using Microsoft Teams, and another four using Zoom Meetings. On average, each interview took 40 - 60 minutes to complete. The interviews were audio-recorded using an external recorder with consent from the participants. The researcher also took notes of key points during the interviews, which helped to facilitate data analysis.

To ensure the credibility and reliability of data collection, a pilot survey and interview were held with a senior staff member of MyCERT. The pilot survey and semi-structured interview were conducted on 14 July 2020, with the survey first followed by the interview. The pilot survey and interview also ensured the feasibility and appropriateness of the survey and interview questions and helped improve the questions based on the feedback received during the pilot. The pilot study also prepared and ensured the interviewer's readiness for the interview. It should be noted that the results from the pilot survey and interviews were not used for the study's data analysis.

### Recruitment of Participants

The study required specific knowledge, understanding and experiences of national CSIRTs' real-world operations to supplement the necessary information to answer the RQ. Therefore, the selection of participants for the study was "purposive" instead of "random" and intentionally selected staff members of national CSIRTs to participate in the study [55]. This was essential to gain accurate, meaningful, and rich insights to answer the research questions [55]. Feedback from staff members of CSIRTs is vital to get insights into the operations of CSIRTs as they have significant experiences in incident responses, much needed when intending to improve CSIRT practices [6].

The researcher's contacts and professional networking with MyCERT and several national CSIRTs allowed easy recruitment of participants for the study. In the first phase, 13 participants from MyCERT were recruited via a gatekeeper. The purpose is to focus the study on the operational practices of a single national CSIRT. Thirteen

[15] See https://web.whatsapp.com/

[16] See https://www.microsoft.com/en-ww/microsoft-teams/

[17] See https://www.zoom.us/

[18] See https://meet.google.com/

Table 9: List of National CSIRTs and their Breakdown into the Two Phases and Data Collection Methods

| Phase | National CSIRT | Website | No. of Participants | |
|---|---|---|---|---|
| | | | Survey | Interviews |
| 1 | MyCERT (Malaysia) | `https://www.mycert.org.my/` | 13 | 7 |
| 2 | CERT.at (Austria) | `https://www.cert.at/` | 1 | 1 |
| | BGD e-GOV CIRT (Bangladesh) | `https://www.cirt.gov.bd/` | 1 | 0 |
| | CSIRT-RD (Dominican Republic) | `https://cncs.gob.do/` | 1 | 0 |
| | CERT-FR (France) | `https://www.cert.ssi.gouv.fr/` | 0 | 1 |
| | JPCERT/CC (Japan) | `https://www.jpcert.or.jp/` | 1 | 0 |
| | CERT-PH (Philippine) | `https://www.ncert.gov.ph/` | 1 | 1 |
| | Sri Lanka CERT/CC | `https://www.cert.gov.lk/` | 1 | 1 |
| | INCIBE-CERT (Spain) | `https://www.incibe-cert.es/` | 1 | 0 |
| | SWITCH-CERT (Switzerland) | `https://www.switch.ch/` | 1 | 1 |
| | TwCERT/CC (Taiwan) | `https://www.twcert.org.tw/` | 1 | 0 |
| | US-CERT (USA) | `https://us-cert.cisa.gov/` | 1 | 0 |
| | One anonymised national CSIRT | | 2 | 0 |
| | | **Total** | 25 | 12 |

participants from MyCERT willingly agreed to participate in the online survey, while seven agreed to participate in a follow-up semi-structured interview.

In the second phase, 13 participants from 12 other national CSIRTs were recruited through the researcher's contact with the CERT Division of the Software Engineering Institute (SEI) of Carnegie Mellon University (CMU), USA [232]. The CERT Division of CMU helped broadcast this study in the national CSIRTs' mailing list. Twelve participants willingly agreed to participate in the online survey, while five agreed to participate in a follow-up semi-structured interview. Notably, two participants from one anonymised national CSIRT participated in the online survey, while other national CSIRTs were represented by one participant each.

A copy of the recruitment email used for the study is available at (`https://cyber.kent.ac.uk/research/CSIRTs/Interview-Survey/Recruitment-Email.pdf`).

A summary of national CSIRTs who participated in the study is reflected in Table 9.

### 4.2.2 Data Analysis

In this study, the data collected from the online survey questionnaires consist of open-ended texts (for instance, to name the data and tools participants use in the operations) and close-ended texts (for example, binary responses, such as "YES/NO", answers to a set of multiple choice questions and participants' level of agreement towards a statement or question – from strongly agree to strongly disagree. Hence, a *quantitative approach* with descriptive statistics, explained in Section 1.4.2 was used to analyse the survey data of close-ended questionnaires, to describe and summarise the data into numbers and graphs [66]. The survey questionnaires with open-ended texts summarised the data into a table as shown in Table 10.

A *qualitative approach* was primarily used to explore and generate main concepts and themes from the semi-structured interview data and put them into the context of the research question [276]. Moreover, a qualitative approach is suitable when research literature on a particular topic is sparse, as reflected by this study's topic. This allowed the exploration of as many insights from the interview data as possible.

The semi-structured interview data were analysed using a thematic analysis, explained in Section 1.4.2. Thematic analysis aligns with this scope to capture new themes and some patterned meanings running across the data. This allows interpretation of the data, aligning with the research question of this study [33]. In the thematic analysis, a manifest approach was primarily adopted to describe what is directly observed without assigning a deeper interpretation or deeper meaning during analysis. A minor latent approach was adopted to interpret the underlying meaning of the data minimally, focusing on the implied value of the data from the researcher's judgement. The themes identified in the data will then inform the research gaps and future needs of national CSIRTs' operations [33]. Content analysis was not adopted to analyse the interview data in this study as it does not fit the scope of the study, which is to identify themes and patterns in the data. Content analysis suits a study that seeks to categorise and quantify qualitative data more objectively.

A coding procedure was adopted to analyse the semi-structured interview data to generate themes [33], described in the next paragraph.

### Coding of Data

Coding was used in this study to find significant information in the data [22; 98; 33]. This study adopted an inductive coding approach to coding the semi-structured interview data. The coding type used is descriptive coding, explained in Section 1.4.2, which helps to capture themes or patterns from the interview data [33]. This aligns with thematic analysis, which focuses on identifying descriptive codes required to answer the RQ. Memoing, a helpful step during the coding process [1], containing the researcher's observations and reflections about the data, was adopted during the coding.

Before coding, the following process was followed first in preparing the interview data for subsequent coding:

1. Transcribed the audio recordings of the 12 semi-structured interview transcripts manually, without transcription software – for privacy and accuracy purposes. The researcher listened to the audio recordings several times for accurate transcription of the recordings and also used "member checking" to ensure the accuracy of the transcribed transcripts [45].

2. Labelled all transcripts with names of the national CSIRTs, e.g. MyCERT, CERT.at. If there is more than one interviewee from the same national CSIRT, the transcript is labelled as P1-MyCERT, P2-MyCERT, or P3-MyCERT; P stands for Participant. The interviewees' names were not used for privacy purposes.

3. Loaded all transcripts into a qualitative data analysis (QDA) software, Atlas.ti version 8.4.5, one of the most widely used QDA software among researchers [147]. Using a QDA software tool saves time and makes data analysis more manageable, ensuring credibility and transparency of analysis [100].

After the data preparation, the next step is coding the data, as in the below process:

1. Read the *transcripts* in two thoughts: a) read for the first time to get an overall understanding and impression of the data, and b) read several times to gain a thorough understanding of the data, detailed insights and a better sense of the data.

2. Coding of the *transcripts* by adopting a systematic step-by-step coding process [218; 1]. The coding task was made easy with a QDA software [147]. The coding began with highlighting *segments* of the transcripts of particular interest for the study to answer the research questions.

3. Generated *codes* through inductive coding of the highlighted segments of the transcripts. The coding involved an iterative process, constantly moving back and forth between the interview data and the codes [33]. The codes described the data that captured the interviewees' views or opinions without the researcher putting any own thoughts on the codes while keeping the RQ in mind [1; 33].

4. Grouped the codes, using a bottom-up approach, by appraising the codes to determine which codes belong to a similar group, thereby forming a *category.* Codes are also grouped based on their semantic meanings. A category consists of codes dealing with the same issue or similar substances. These are often short, factual sounding, with manifest meaning visible in the data and with limited interpretation by the researchers [61]. These categories derived from the coding process form the *themes* of the study.

5. Present report and findings – the *themes* formed from the code categories that answer the research question.

## 4.3 Results

This section presents the survey and semi-structured interview results. The results from the survey will be presented first, followed by the semi-structured interviews.

### 4.3.1 Results from the Online Survey

The results from the online survey are primarily based on descriptive statistics analysis of quantitative data from closed-ended texts of the survey questionnaire, reflected in Figures 8–11. This is then combined with a minor qualitative synthesis of data from open-ended texts of the survey questionnaire, such as the names of the tools, public data and closed-source data used in the participating national CSIRTs and their purposes. Minor qualitative synthesis combines qualitative information from open-ended texts with quantitative data from closed-ended texts of the survey questionnaire.

#### Demographic Information about Participants

Thirteen participants from MyCERT and 12 participants from other national CSIRTs participated in the online survey, shown in Table 9. The participants' demographic consists of Team Leaders (one from MyCERT, three from other national CSIRTs), Executives (none from MyCERT, two from other national CSIRTs) and Analysts (12 from MyCERT, seven from other national CSIRTs) illustrated in Figure 8. Notably, Team Leaders lead incident response teams, Analysts analyse or investigate incidents while Executives support the operations of national CSIRTs. Additionally, all participants demonstrated some experience working at national CSIRTs, shown in Figure 9, with most of them between one and five years of working experience.

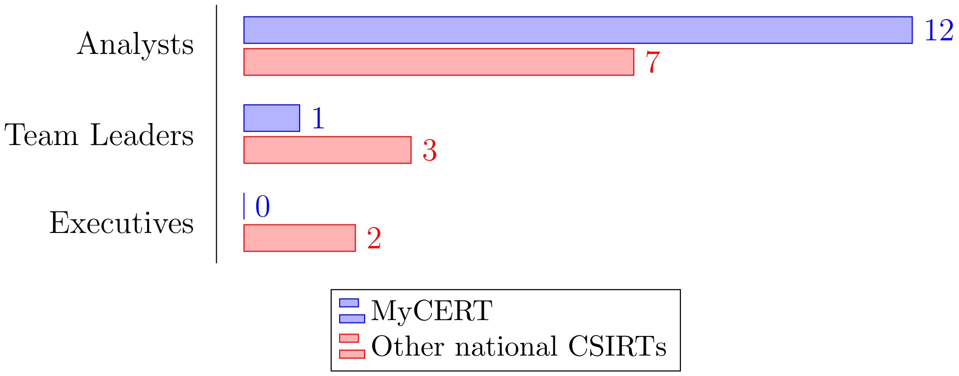


Figure 8: Survey Participants' Roles at their National CSIRTs

#### Methods of Incident Investigation

The survey asked participants how incidents are investigated at their national CSIRTs and found that seven from MyCERT and four from other national CSIRTs use semi-automated methods to investigate incidents. Three from MyCERT and one from other national CSIRTs use a manual approach. One from MyCERT and three from other national CSIRTs use a combination of manual and semi-automated approaches. Two from MyCERT use a combination of semi-automated and automated, one from other national

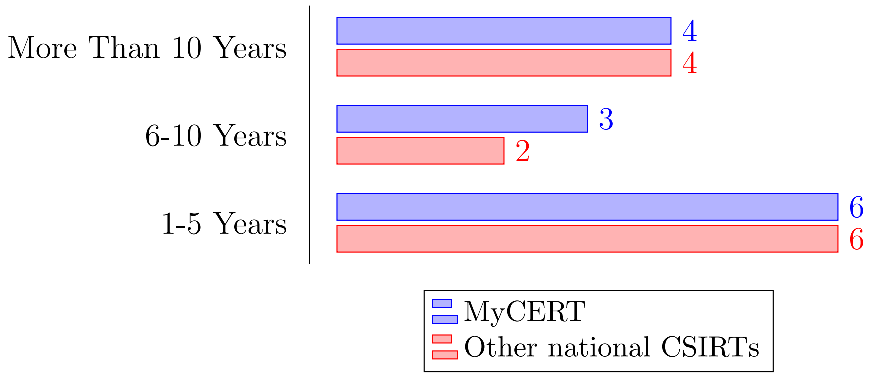


Figure 9: Survey Participants' Working Experience at their National CSIRTs

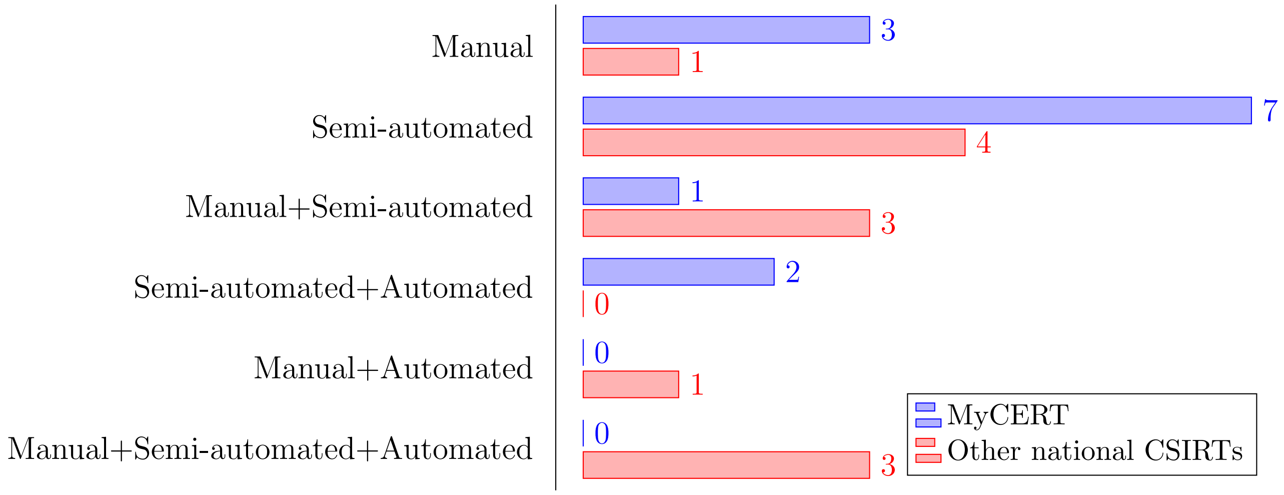


Figure 10: Method of Investigating Incidents Reported by Participants in the Survey

CSIRTs uses a combination of manual and automatic, and three from other national CSIRTs use a combination of manual, semi-automated and automated. A summary of the methods used for investigating incidents in the surveyed national CSIRTs is shown in Figure 10. Notably, none of the participants surveyed used a fully automated method, but this option is not shown in the above Figure.

### Participants' Experience with Public Data

The survey result showed that 13 participants from MyCERT and 12 from other national CSIRTs had used public data for incident investigations[19] .

Participants from MyCERT and other national CSIRTs use various public data to facilitate incident response. These consist of publicly shared malware samples such as those from Virus Total (`https://www.virustotal.com/`), publicly disclosed malicious IP addresses and URLs of phishing websites, public DNS records, data from search engines

[19] Two participants chose 'no' when being asked about their experience with public data but answered 'yes' to another question about validating public data. One of them participated in an interview and clarified that they did use public data. Another participant was not interviewed, but it is believed they also used public data

such as Google Search, public information from other national CSIRTs, public data from honeypots, ad hoc public data feeds and datasets such as Shodan (`https://www.shodan.io/`), Google Hacking Database (`https://www.exploit-db.com/google-hacking-database`) and Censys (`https://censys.io/`), publicly released threat reports by security organisations, public data on online social media (e.g., Twitter), and news reports. The participants also mentioned public Application Programming Interfaces (APIs), sensors deployed in multiple network gateways and domain registry databases, and the dark web without explaining what they referred to exactly, but presumed to be public data sources, based on the researcher's understanding.

Notably, all participants from MyCERT (13) and other national CSIRTs (12) mentioned validating public data before using them for investigation[20] . Participants used various tools and methods for validation, including ad-hoc experiments[21] , cross-checking the data with trusted organisations such as national CSIRTs and independent cyber security experts and other researchers, and cross-checking the data with data collected from different platforms. Based on the survey result, no systematic procedures exist to validate the public data for reliability and authenticity in the surveyed national CSIRTs.

### Participants' Experience with Closed-Source Data

To put the use of public data into the proper context, the survey also had a question about the use of closed-source data within national CSIRTs. All the surveyed participants except one from another national CSIRT stated using closed-source data for incident responses. These data are generally from victims and organisations reporting incidents and sharing information about cyber threats with national CSIRTs. Such closed-source data include artefacts from compromised systems such as system logs, malware samples, digital forensic images, email headers, URLs of phishing websites, malicious IP addresses and domain names, defaced web pages, and closed-source intelligence about threats.

All participants from MyCERT (13) and other national CSIRTs (12) also expressed their difficulties in analysing closed-source data. This is because closed-source data often lack details about incidents such as the date and time stamp, lack contexts such as the chronology or background of an incident, incompatible data formats and unstructured data that require bespoke parsing tools for analysis[22] , and extensive data that requires human experts to interpret the data.

[20] One participant reported that he did not validate public data, but during a follow-up interview with him, it was clarified that he made a mistake for this question in the survey.

[21] Two MyCERT participants used the term "proof of concept" (POC) for this method. Judging on the researcher's knowledge as an employee of MyCERT, POC is an experiment whereby staff run a malware sample in a sandbox environment to validate it being malware).

[22] Although structured data with an unknown or unsupported format also require bespoke parsing tools, this challenge was not mentioned by any participants.

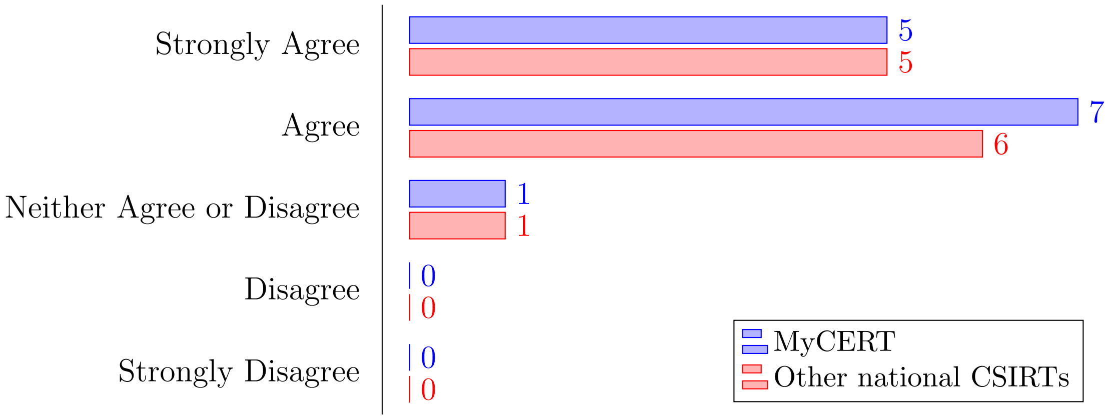


Figure 11: Survey Participants' Perceptions on the Usefulness of Public Data and OSINT Tools

### Participants' Experience with Software Tools

The survey also asked about the different software tools (including online services) participants use for incident response. Participants mentioned using OSINT and non-OSINT tools. Notably, the survey found most OSINT tools used by the participants are free, whilst very few are commercial and used for particular areas of investigation, such as digital forensics. Free tools mentioned by participants are listed in Table 10. Participants said that the selection of tools in the national CSIRTs largely depends on the complexity of the incidents, staff members' role in the operations (i.e. first-level support or second-level support), and staff expertise in using tools. A few participants were reluctant to disclose the tools they use in the operations, especially commercial ones, due to the worry of over-disclosing operational practices of their national CSIRTs. The survey also found some tools were considered public data sources since their functionalities include or are about providing public data (e.g., Google Hacking Database).

### How Participants Perceived the Usefulness of OSINT Tools and Public Data

The survey asked participants how they perceived the usefulness of OSINT tools and public data to support incident response in the operations of national CSIRTs on a five-point Likert scale. All except one participant from MyCERT (12) agreed and strongly agreed that OSINT tools and public data are useful for national CSIRTs. The same was observed for participants from other national CSIRTs, whereby 11 out of 12 agreed or strongly agreed on the usefulness of public data and OSINT tools for national CSIRTs. The detailed breakdown of participants' perceptions is shown in Figure 11.

On combining public data and closed-source data to facilitate the incident investigation, a majority (8 out of 13, 62%) of participants from MyCERT agreed or strongly agreed that they often get better investigation results when combining public data with closed-source data. Similarly, participants from other national CSIRTs (9 out of 12,

Table 10: Tools Mentioned by Survey Participants (OS = fully open-source, FW = freeware without open source, FW+OS = freeware with partial open source, FOS = free online service)

| **Tool** | **URL** | **Brief description** |
|---|---|---|
| VirusTotal | `https://www.virustotal.com/` | OSINT, FOS, malware analysis |
| Google Hacking Database | `https://www.exploit-db.com/google-hacking-database` | OSINT, FOS, information gathering technique for advanced Google searching. |
| AlienVault OTX | `https://otx.alienvault.com/` | OSINT, FOS, online threat indicators such as IOCs |
| MISP | `https://www.misp-project.org/` | OSINT, OS, for obtaining, sharing and corelating IOCs |
| Maltego Community Edition | `https://www.maltego.com/` | OSINT, FW, for profiling threat actors |
| Shodan | `https://www.shodan.io/` | OSINT, FOS, for searching into the Internet of Things |
| OSINT Framework | `https://osintframework.com/` | OSINT, OS, FOS, online collector of information from free tools and resources |
| IntelMQ | `https://intelmq.readthedocs.io/` | OSINT, OS, for automatic collecting and processing security feeds |
| Taranis | `https://github.com/NCSC-NL/taranis3` | OSINT, OS, for monitoring and analysing news items and writing security advisories |
| Censys Search | `https://search.censys.io/` | OSINT, FOS, for discovering, monitoring, and analysing Internet-connected devices |
| Hybrid-Analysis | `https://www.hybrid-analysis.com/` | FOS, a platform for analysing malware (in-depth static and dynamic analysis) |
| REMnux | `https://remnux.org/` | FW, malware analysis |
| Windows Process Monitor | `https://docs.microsoft.com/en-us/sysinternals/downloads/procmon` | FW, to monitor running processes on Windows |
| MobSF | `https://github.com/MobSF/Mobile-Security-Framework-MobSF` | OS, for mobile device penetration testing and malware analysis |
| Wireshark | `https://www.wireshark.org//` | OS, network protocol analyser |
| Security Onion | `https://securityonionsolutions.com/` | OS, network protocol analyser |
| Windows Sysinternal Suite | `https://docs.microsoft.com/en-us/sysinternals/downloads/sysinternals-suite` | FW, a Windows toolset for troubleshooting |
| Joe Sandbox | `https://www.joesandbox.com/` | OS, FOS, detects and analyses potential malicious files and URLs on Windows, Android, Mac OS, Linux, and iOS for suspicious activities. |
| PeiD (PE iDentifier) | `https://www.aldeid.com/wiki/PEiD` | FW, PE file analysis for detecting packers, cryptors and compilers used |
| Hxd | `https://mh-nexus.de/` | FW+OS, hex editor |
| SiLK (System for Internet-Level Knowledge) | `https://tools.netsa.cert.org/silk/` | OS, network flow collection and storage infrastructure |
| nfdump | `https://github.com/phaag/nfdump` | OS, for collecting and processing netflow and sflow data |
| Nfsen | `http://nfsen.sourceforge.net/` | OS, graphical frontend of nfdump |
| DomainTools WHOIS | `https://whois.domaintools.com/` | FOS, domain name and IP address whois lookup |
| Cuckoo Sandbox | `https://cuckoosandbox.org/` | OS, malware analysis sandbox |
| Apache Pulsar | `https://pulsar.apache.org/` | OS, for netflow analysis |
| Apache Flink | `https://flink.apache.org/` | OS, for netflow analysis |

75%) agreed or strongly agreed that they often get better investigation results when they combine public data with closed-source data. The detailed breakdown of participants' perceptions is shown in Figure 12.

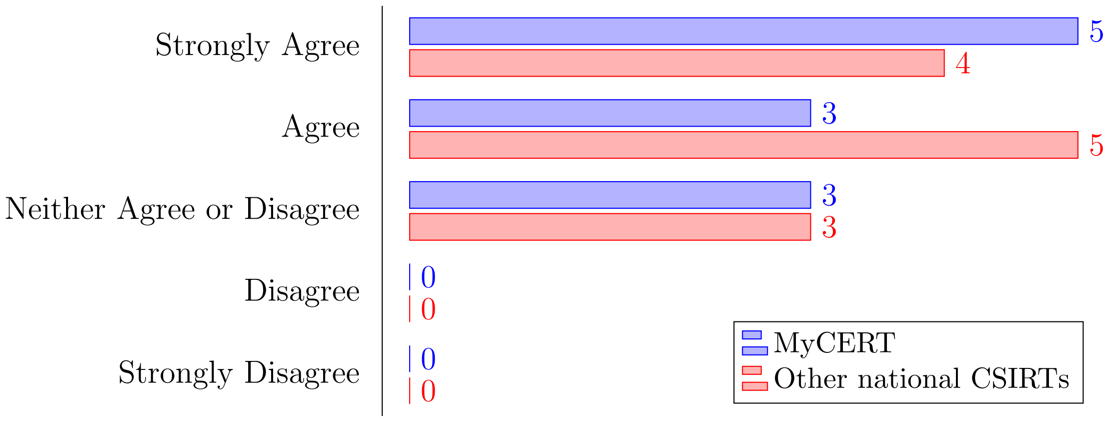


Figure 12: Survey Participants' Perceptions if Combining Public Data and Closed-source Data Provides Better Results

### 4.3.2 Results from the Semi-structured Interviews

The online survey led to some significant findings, such as all (25) participants having used public data to facilitate incident response and most participants (23) having perceived OSINT tools and public data as useful for the operations of national CSIRTs. The results from the semi-structured interviews further consolidated and substantiated findings from the online survey.

Seven participants from MyCERT and five other national CSIRTs participated in the online semi-structured interviews, shown in Table 9. The 12 semi-structured interview transcripts were analysed using a thematic analysis method to generate codes that form code categories and then into "themes" to answer the RQ. The coding flow of the 12 semi-structured interview transcripts is described in Section 4.2.2. A total of 179 quotations (segments) were extracted from the 12 semi-structured interview transcripts. Of the 179 quotations (segments), 256 codes were extracted and grouped into four code categories. The four code categories were reviewed before they were formed into themes. All interviewees were asked to give general background information about their national CSIRTs. Such information was not encoded as it does not relate to the RQ but was summarised directly from the interview transcripts to provide the study context. The code book for the study is available at (`https://cyber.kent.ac.uk/research/CSIRTs/Interview-Survey/Codebook.pdf`).

The study identified four themes, presented in the below order:

1. How the interviewed national CSIRTs use OSINT, free tools and public data in the operational practices.

2. How the interviewed national CSIRTs' staff perceive the usefulness of OSINT, free tools and public data for the operational practices.

3. Operational challenges faced by the interviewed national CSIRTs.

4. Reporting and sharing information about cyber incidents in the interviewed national CSIRTs' operations.

The general background (from the early discussions with the interviewees) is presented to provide the study context, followed by the themes.

### General Information about National CSIRTs

Initially, all interviewees were asked to introduce their national CSIRTs' operations and their experiences with the national CSIRTs.

Most of the interviewed national CSIRTs manage security incidents with the help of a help desk system or a ticketing system, such as the Request Tracker Incident Response (RTIR) and Online Ticket Request System (OTRS) – open-source systems [169]. Meanwhile, standard operating procedures (SOP) guide the interviewed national CSIRTs in responding to incidents. The content of the SOPs is primarily referenced from the National Institute of Standards and Technology (NIST) Computer Security Incident Handling Guide [42] and the SANS Institute's Six Steps of Incident Handling [139]. Interviewees from MyCERT explained that there are three tiers in MyCERT operation, Tier 1, 2 and 3, each responsible for specific tasks. Similarly, the interviewee from the SWITCH-CERT stated that a dedicated tier system is established in SWITCH-CERT to handle only complex malware incidents. A participant from the CERT-FR reported having different dedicated teams focusing on the broader areas of operations, such as dedicated teams to handle data acquisition and analysis work. Interviewees from other than MyCERT, SWITCH-CERT and CERT-FR do not have tier systems in their operations. Nevertheless, the interviewees expressed their interest and planned to implement a tier-based operation in their national CSIRTs.

Interviewees were also asked to explain how incidents are classified in the operations. In general, CSIRTs use a pre-defined incident classification scheme [4; 157; 97]; hence, the study is interested in gaining more insights into this question. The study revealed that a standardised and universal incident classification scheme is currently unavailable for the interviewed national CSIRTs. A small follow-up study was undertaken in this research to gain insights into incident classifications in national CSIRTs. A similar finding was revealed, and a list of different classification schemes was identified as being used by national CSIRTs worldwide [169].

All interviewees from MyCERT mentioned handling *technical incidents* (such as malicious codes, spams, vulnerability handling, intrusions and intrusion attempts, and denial of service attacks) and *non-technical incidents* (such as harassment, content-related, and fraud). Three other national CSIRTs handle technical incidents only, while another two handle technical and non-technical incidents, similar to MyCERT.

The basic information about national CSIRTs' operations provides some context to understand better the themes that emerged from the semi-structured interview data described in the following section.

### Theme 1: How the interviewed national CSIRTs use OSINT, free tools and public data in the operational practices

This theme concerns how OSINT, free tools and public data are used in the interviewed national CSIRTs. The semi-structured interviews were different from the online survey, where questions about free tools were not explicitly asked in the survey to avoid unnecessary confusion. Nevertheless, the subtle relationships between OSINT tools and free tools were covered in the semi-structured interviews, as the survey found that most OSINT tools used in the participating national CSIRTs are free tools.

Considering the online survey results, it is unsurprising that all seven interviewees from MyCERT confirmed using public data in their daily incident response work. They mentioned that public data are used primarily to obtain richer (contextual) information and detailed insights or attributes about a particular incident. Apart from facilitating investigations, one interviewee from MyCERT mentioned that public data is also used as input for the production of security advisories and alerts regarding new threats- as part of national CSIRTs' information sharing on current threats for the general public.

The tools mentioned by MyCERT participants during the online survey were confirmed during the online semi-structured interviews. Additional tools were also mentioned during the interviews, as shown in Table 11, such as IDA Pro and Splunk free version to support incident responses, precisely for malware and log analysis. On using OSINT tools just over half of the interviewees from MyCERT (4 out of 7) use OSINT tools in their incident response work, while all interviewees use free tools (inclusive open-source tools). It should be noted that free tools described in Theme 1 have no reference to OSINT tools.

The findings from MyCERT were compared with those of other national CSIRTs. All five interviewees from other national CSIRTs use public data, similar to MyCERT. The interview also found all five interviewees from other national CSIRTs use OSINT tools, unlike MyCERT. It is also common for all interviewees to use free tools in the operations, similar to MyCERT. The purpose of using OSINT, free tools, and public data is consistent with MyCERT, which is for incident responses, investigation of incidents, and production of security advisories and alerts.

Most tools mentioned by other national CSIRTs during the online survey were confirmed during the online semi-structured interviews. For instance, IntelMQ listed in Table 10 was mentioned by interviewees from CERT.at, SWITCH-CERT and CERT-FR. Some interviewees also mentioned several free tools not mentioned in the online survey, compiled in Table 11. One interviewee also informed about tools developed by CERT.at available at the GitHub portal (https://github.com/certat), with many free and open-source tools listed.

Table 11: Additional Tools Mentioned in the Interviews (OS, FW, FW+OS: see the caption of Table 10)

| Tool | URL | Brief description |
|---|---|---|
| IDA Pro Free Version | https://hex-rays.com/ida-free/ | FW, malware analysis |
| Splunk Free Version[23] | https://www.splunk.com/en_us/download.html | FW, for log analysis |
| Notepad++ | https://notepad-plus-plus.org/ | OS, universal file editor for log analysis |
| Kali Linux | https://www.kali.org/ | FW+OS, security-enhanced Linux distribution with many useful tools |
| Nmap | https://nmap.org/ | OS, for network discovery and security auditing |
| OpenCTI | https://www.opencti.io/ | OS, for storing, organising, visualising and sharing knowledge on cyber threats |
| Tiny Tiny RSS | https://tt-rss.org/ | OS, web-based news feed reader and aggregator |

The semi-structured interviews found that OSINT, free tools and public data were used ad hoc at the interviewed national CSIRTs without referring to any specific procedures. The following two interviewees mention this when asked about any procedures governing the use of public data in the interviewed national CSIRTs:

> "*No, it's done on an ad hoc basis.*" (Interviewee, CERT.at)

> "*Unfortunately, we don't have a procedure. Not yet. Actually, we're slowly building the procedure, this one we don't have yet.*" (Interviewee, CERT-PH)

### Theme 2: How the interviewed national CSIRTs' staff perceive the usefulness of OSINT, free tools and public data for the operational practices

This theme concerns how interviewees perceive the usefulness of OSINT, free tools and public data for national CSIRTs' operations. All interviewees from MyCERT (7) and other national CSIRTs (5) perceived OSINT, free tools, and public data as useful for national CSIRT. They all perceived public data as useful for better comprehending cyber incidents from different perspectives. According to the interviewees, public data also

[23] The free version was available in the past (see https://www.splunk.com/view/SP-CAAAE66), which is not available any longer (only free trials for most products from Splunk).

serves as a pivot point to other useful information about incidents, as mentioned by one interviewee from MyCERT:

> "*We can get different views of the data given to us, get different points of view from different angles, different additional views, either new information from other data or correlation.*" (Interviewee, MyCERT)

Interviewees from MyCERT and other national CSIRTs also perceived public data is more often up-to-date (an essential requirement for data to ensure timely responses to incidents). One interviewee from MyCERT mentioned that public data could help save time because crucial information, such as analysis of new malware variants, is more often available in public data. Such information can be consumed directly to augment an ongoing investigation without conducting a further analysis. Echoing the online survey result of this study, all interviewees from MyCERT and other national CSIRTs confirmed that public data often provides more meaningful and rich investigation results when used together with closed-source data.

Five interviewees from MyCERT and five from other national CSIRTs perceived free tools as useful for national CSIRTs' operations. To reinforce this perception, one interviewee from MyCERT mentioned that free tools could produce results comparable to those of commercial tools:

> "*We use IDA Pro, the free version, to check the network behaviour of the binary and to see what the malicious binary is doing towards the operating system (OS). We still can get results of the static analysis even if it is free, maybe some features are limited. Other people use [ANONYMISED] (a commercial tool). The result is [the] same, just that it is presented in a different format.*" (Interviewee, MyCERT)

Another interviewee from Sri Lanka CERT/CC mentioned how free tools helped them to successfully identify a command & control (C&C) server involved in malicious activities. This allowed them to inform the relevant Law Enforcement Agency (LEA) to remove the identified C&C server.

Interviewees perceived a key advantage of free tools is the availability of such tools freely on the Internet while open-source tools have advantages in customisation besides an active community around open-source tools. A participant from the SWITCH-CERT stated this:

> "*The benefits are free tools are free, and if you have open-source tools, you add or modify the source codes as you like, as intended. If you have open source in the community, the community develop further for the community.*" (Interviewee, SWITCH-CERT)

Despite the positivity, interviewees also highlighted issues and problems with public data, OSINT, and free tools, described in the section below. These are primarily related to reliability, authenticity, usability, and lack of validated OSINT and free tools, indicating a gap.

### Theme 3: Operational challenges faced by the interviewed national CSIRTs

Interviewees also highlighted issues and problems with OSINT, free tools, and public data, as described below. These are primarily related to reliability, authenticity, usability, and lack of validated OSINT and free tools, indicating a gap.

The interviewees were asked about the operational challenges at their national CSIRTs. Such challenges can be fragmented into three key areas: (1) public data, (2) OSINT and free tools and (3) resources – as summarised in Table 12, which are largely self-explanatory. Some interviewees also mentioned less relevant challenges and challenges based on inaccurate information. Such challenges include some issues with closed-source data, insufficient budget to purchase commercial tools and the challenge based on an erroneous claim of an open-source tool. Such irrelevant challenges and problematic "claims" are excluded from the study. Some challenges represent isolated or very subjective opinions of one or just a few interviewees, which were also excluded. Such selective synthesis of interviewees' feedback ensured the challenges reported in this section were more common and representative of the interviewed national CSIRTs.

Six interviewees (four from MyCERT and two from other national CSIRTs) expressed challenges concerning OSINT and free tools. According to the interviewees, the free version of such tools comes with limited features compared to the commercial version. Another challenge with OSINT and free tools is that such tools are often outdated and take a long time to update. Moreover, most free and open-source tools cannot process extensive data. A related challenge is when acquiring commercial tools would require some budget, and a few interviewees (3 out of 12) considered budget a challenge for national CSIRTs.

A challenge highlighted by seven interviewees (five from MyCERT and two from other national CSIRTS) concerns resources. According to the interviewees, this consists of an insufficient and incompetent workforce (staff), such as a lack of skilled staff in using OSINT tools to do analytic work.

The challenge concerning public data was raised by eight interviewees (five from MyCERT and three from other national CSIRTs). This includes dealing with overwhelming and verbose public data. The reliability of such data is not confirmed; these are primarily unstructured and require special skills and expertise to conduct analysis, as commented below:

> "*We've sometimes too much data, plenty of data, sometimes reliable and*

Table 12: Operational Challenges with Public Data, OSINT and Free Tools (Synthesised from Survey and Interviews)

| Area | Number of Interviewees | | | Challenge |
|---|---|---|---|---|
| | MyCERT | Other National CSIRTs | Total | |
| **Public Data** | 5 | 3 | 8 | Lack of validated public data that can be considered reliable |
| | | | | The huge amount of (public) data making it difficult to find useful information |
| | | | | Lack of sufficient data shared by some organisations (e.g., some users of MISP do not actively share data) |
| **OSINT+Free Tools** | 4 | 2 | 6 | Lack of officially validated tools that can be used by CSIRTs |
| | | | | Usability issues of some tools (e.g., limited or broken features) |
| | | | | Lack of enough tools to process unstructured data |
| | | | | Lack of tools that can process big data at high speed (e.g., for real-time netflow analysis) |
| **Resources** | 5 | 2 | 7 | Insufficient workforce (e.g., due to loss of competent staff) |
| | | | | Insufficient skills and expertise (e.g., on OSINT tools and data analytics) |

> *sometimes not, lack of resources, skills and expertise in analysing data*" (Interviewee, CERT-FR)

.

One interviewee from another national CSIRT voiced the challenge in prosecuting perpetrators, as expressed below:

> "*Sometimes we found the IP address of a particular person, but challenge is*

*to prosecute successfully*" (Interviewee, Sri Lanka CERT/CC)

.

According to the interviewee, the challenge arises when national CSIRTs are entrusted with obtaining the necessary technical evidence that is admissible in courts for prosecution. The interviewee also suggested establishing a good rapport with regulatory bodies, LEAs, and Internet service providers (ISPs), which may help overcome such a challenge in national CSIRTs.

**Theme 4: Reporting and sharing of information about cyber incidents in the interviewed national CSIRTs' operations**

The semi-structured interviews revealed a trend in incident reporting and information sharing between non-CSIRT organisations (including victims) and the interviewed national CSIRTs[24] , and the reasons behind such a trend. All seven interviewees from MyCERT revealed that not all victims and organisations experiencing incidents report the incident to MyCERT. The interviewees mentioned several reasons why organisations are not reporting incidents to MyCERT. Six interviewees from MyCERT expressed concerns over brand reputation due to the publicity of the incidents if reported to a third party, such as a national CSIRT.

The following five reasons (each by one interviewee from MyCERT) deter victims and organisations from reporting incidents to respective national CSIRTs:

1. Some organisations have stringent internal policies on data privacy and information disclosure. This is consistent with a previous study that found companies refrain from reporting incidents due to data protection and company policies [19].

2. Some victims and organisations believed they could handle the incidents themselves.

3. Some organisations hire third-party cyber security experts to handle the incidents.

4. Victims, e.g. home users, do not report incidents due to privacy concerns and other personal reasons.

5. The data needed as evidence to support incident reporting may not be available. Interestingly, one interviewee mentioned that some organisations report incidents without details or evidence of the incident. This is because such organisations are only interested in obtaining more information about an incident rather than seeking assistance from national CSIRTs to handle the incident.

[24] There is active information sharing among national CSIRTs using free tools to facilitate such sharing, e.g., MISP and IntelMQ. However, this aspect was not discussed during the interviews.

The trend in incident reporting and information sharing revealed by interviewees from MyCERT was consistent with other interviewed national CSIRTs. Four interviewees from other national CSIRTs mentioned difficulties getting victims and organisations to report incidents to the respective national CSIRTs. An interviewee from CERT.at mentioned that Austria has a regulation that mandates cyber incident reporting in Austria[25] . Some non-EU countries have also passed similar legislation, e.g., in the USA [42]. The USA is also introducing legislation – Cyber Incident Notification Act of 2021 [13], but this has not been passed yet. The Bill was introduced in the Senate (07/21/2021). This Bill requires federal agencies and certain entities to report cyber security intrusion incidents to the Cybersecurity and Infrastructure Security Agency (CISA) and addresses related issues. In March 2022, President Biden signed a law, the Cyber Incident Reporting for Critical Infrastructure Act of 2022 (CIRCIA) [49]. The law requires the CISA to develop and implement regulations to ensure cyber incidents and ransomware payment incidents are reported to the CISA. Such incidents will allow CISA to deploy resources and assist victims suffering attacks rapidly, analyse incoming reporting across sectors to spot trends, and quickly share that information with network defenders to warn other potential victims.

Interviewees from other national CSIRTs mentioned several reasons for the lack of incident reporting among victims and organisations experiencing incidents. These reasons are different from those mentioned by MyCERT interviewees. One interviewee from CERT-PH (Philippines's national CSIRT) stated two primary reasons for the lack of incident reporting: 1) the lack of trust, especially from the private sector, in the national CSIRT and 2) the lack of knowledge about the national CSIRT (CERT-PH was in service for only three years at the time of the interview). Another reason mentioned by an interviewee from SWITCH-CERT is because victims may be afraid of being blamed for the incident if such an incident is reported to relevant authorities[26] :

> "*It is not just the problem of customer trust. It is more like having fear to be blamed if something happens. I think they have fear that someone would blame, that it happened to you, how could it be.*" (Interviewee, SWITCH-CERT)

The result and finding concerning incident reporting is consistent with a previous study, which saw only a tiny percentage (15%) of all cyber attacks reported to authorities [20]. Another study found that organisations are not reporting cyber incidents as they fear negative customer perceptions, potential damage to brand reputation and loss of trust by the public [185], consistent with this study.

[25] The interviewee referred to the NISV (Netz- und Informationssystemsicherheitsverordnung `https://www.ris.bka.gv.at/Dokumente/Erv/ERV_2019_2_215/ERV_2019_2_215.html`), the national law defined according to the EU NIS Directive. The law does not mandate compulsory reporting of cyber incidents to CERT.at, but just for incidents with an "impact on economic and societal activities" that occur to essential and digital services. All EU member states now have such legislation.

[26] This echoes similar findings from previous research on the blame culture in other contexts [124; 80].

## 4.4 Summary

This chapter presented an empirical study with a single national CSIRT followed by 12 national CSIRTs worldwide to understand better the real-world operational practices of national CSIRTs regarding the use of OSINT, free tools and public data. Notably, Chapter 3 is more data-driven, while this chapter is empirical and qualitative, giving a real-world picture of how OSINT, free tools and public data are used and perceived in the surveyed and interviewed national CSIRTs. The study was conducted in two phases, first in a single national CSIRT, followed by 12 national CSIRTs. Findings from both phases are synthesised to understand better the real-world operations of national CSIRTs who participated in the study.

Overall, this chapter found that all surveyed and interviewed national CSIRTs (from both phases of the study) use OSINT tools, free tools and public data to facilitate incident responses. However, whether the usage is more often compared to non-free tools and non-public data remains unproven. The nature of free tools and public data is often less validated, but more people are willing to use them because they are free, lowering the quality or quality threshold of free tools and public data. Therefore, this research is interested in finding a systematic approach for evaluating free tools and public data to ensure only qualified tools and data are used in operations. In contrast, commercial and closed-source tools are often evaluated in some way by the company developing them or by a trusted body. There is some level of evaluation for commercial tools, but it is uncertain for the free ones. The information obtained from the study on how public data is validated is less understood, whilst how OSINT and free tools are validated is unknown. Hence, more work is needed to gain insights into how free tools and public data are evaluated in the operations of national CSIRTs and how such practices can be enhanced with systematic procedures. Therefore, an empirical study using focus groups is undertaken to understand better how free tools and public data are evaluated in national CSIRTs and establish a procedure for evaluating free tools and public data – i.e., criteria for evaluating tools and data, presented in the next chapter.

# Chapter 5

# Understanding Free Tools and Public Data Evaluation Practices of National CSIRTs and Constructing Candidate Criteria for a Systematic Evaluation

This Chapter is a follow-up to the previous Chapter 4. This chapter provides insights into a critical operational gap across national CSIRTs concerning the use of free tools and public data, reported and less understood in the previous chapter – a systematic evaluation of free tools and public data to ensure the quality of such tools and data in the operations of national CSIRTs. The study presented in this chapter was performed to answer RQ3 of this research outlined in Section 1.3. The chapter is structured as follows. Section 5.1 introduces the study and original contributions. Section 5.2 explains the methodology adopted for the study, primarily the data collection – focus group discussion and data analysis – content analysis. Section 5.3 provides the results and findings according to the research question, while Section 5.4 summarises the chapter.

## 5.1 Introduction

National CSIRTs have received much attention due to their crucial role in safeguarding national infrastructures from cyber attacks. To my knowledge, limited research has been done concerning using free tools and public data in national CSIRTs' operational practices. It is known that public data, open data, closed-source data, OSINT, and open-source and free tools are frequently used by national CSIRTs to facilitate their incident responses, as reported in Chapter 4. However, the research literature lacks a systematic discussion on how national CSIRTs would ensure the quality of such tools and data they use in the operations. Some national CSIRTs evaluate the tools and the data they use in an ad-hoc and informal manner to ensure the use of appropriate tools and data [168]. This raises concerns about how such tools and data are selected systematically and properly. This is problematic, especially for free tools and public data, which generally do not undergo quality assurance compared to commercial ones. To allow qualified tools and data to be used in operation, it is essential to evaluate such tools and data using systematic procedures or guidelines – a set of criteria for evaluating tools and data [87]. Such an evaluation is critical to avoid problematic tools and data, such as "a lack of security" [28] and to ensure only effective tools and data are used in national CSIRTs' operations [27].

Hence, it would be interesting to examine further the ad-hoc practices in national CSIRTs on tools and data evaluation reported in Chapter 4 and to understand better issues preventing a systematic practice. The researcher's observation indicates public data, OSINT, and free and open-source tools are often not evaluated (or not being assessed systematically) due to a lack of standard procedures, resources and expertise [167]. The insights gained from the present study are essential to inform the development of a systematic procedure or guideline consisting of criteria for evaluating tools and data that can be used in national CSIRT operations.

This study conducted seven online focus group discussions with 20 staff members from 15 national CSIRTs from Asia-Pacific, America, Africa and Europe. The results of the focus group discussions led to five major findings, summarised as follows:

1. Empirical evidence of how tools and data are currently evaluated in the real-world operational practices of the participating national CSIRTs.

2. The results and efforts from the study are translated into a set of candidate criteria that national CSIRTs, CSIRTs and the broader security operations could use to specify the quality of tools and data.

3. New candidate criteria identified from the study could help refine requirements defined in the ISO/IEC 25000 standard "Systems and software Quality Requirements and Evaluation (SQuaRE)" [253] and other software evaluation models mentioned in the literature.

4. Key findings from the study lay solid foundations for future research and development activities for different stakeholders: 1) national CSIRTs (towards improving their operational practices); 2) software developers and vendors (towards developing tools and data that are more aligned with users' needs); and 3) researchers (towards conducting more targeted research, e.g., in developing more advanced machine learning methods for tools and data evaluation).

The rest of this study is organised as follows. Section 5.2 explains the study's methodology, particularly the data collection and analysis strategies. Section 5.3 presents the results of the investigation.

## 5.2 Methodology

This section explains the methodology used for the study, mainly the data collection and analysis. The primary approach used in this study is qualitative. A minor quantitative approach was used to get an overview of the data by translating it into numbers and figures. The quantitative findings inform the qualitative, where the key findings of the overall study are derived. As the study involved human participants, a research ethics application was submitted to the University of Kent's Central Research Ethics Advisory Group (CREAG). This study received a favourable opinion from the University of Kent Ethics Board on 5 October 2021 under reference number CREAG087-09-2021 to conduct the data collection for the study.

An electronic Participant Information Sheet (PIS) was provided to participants, giving details about the study, explaining the scope of the study, their rights to withdraw without providing a reason, what data will be collected and how it will be processed, stored and used. In addition, an electronic Consent Form was provided to each participant to get their consent. All participants willingly gave their consent to participate in the study and agreed to have their direct quotes included in any research publications resulting from this study, with their personal information anonymised. The Participant Information Sheet (PIS) and the Consent Form used in the study can be found at (`https://cyber.kent.ac.uk/research/CSIRTs/Focus-Group/PIS.pdf`) and (`https://cyber.kent.ac.uk/research/CSIRTs/Focus-Group/Consent-Form.pdf`), respectively.

### 5.2.1 Data Collection

*Focus Group Discussion* was adopted as the data collection method in this study to draw out expert knowledge and experiences from staff members of national CSIRTs through collective group discussions [213]. Eventually, this allowed the collection of richer information through active interactions within the focus groups [213]. Notably, all participants have a common background, expertise and experience related to the topic of the study. Focus groups also allowed the collection of information, as rich as possible, from participants through free-flowing group interactions [89] – aligned with the inductive and exploratory nature of the present study.

Focus group discussions are typically conducted face-to-face. However, the advancement of telecommunication technologies (such as online meetings and virtual discussion platforms) allowed researchers to conduct focus group discussions in an online environment [240; 68]. Previous studies have successfully demonstrated the feasibility of online focus group discussions [241; 70; 90], motivating an online environment for the present research.

The Microsoft Teams online platform (`https://www.microsoft.com/en-my/microsoft-teams/`) was used in this study as the researcher's Institution has a site license for security and compliance purposes. Microsoft Teams is also considered a reliable online communication platform. It allows automatic audio recording of the focus group discussions with reasonable sound quality without an additional external audio recorder. Furthermore, all participants preferred Microsoft Teams as a preferred platform for the study. The interviews were audio recorded, as consented to by participants and transcribed.

The instrument used for the focus group discussions was a *Focus Group Schedule*. It contains the focus group discussion agenda with a list of questions. The questions are arranged in order, from general to specific [141]. The focus group schedule used for the study can be found at `https://cyber.kent.ac.uk/research/CSIRTs/Focus-Group/Focus-Group-Schedule.pdf`. The questions in the Focus Group Schedule are divided into four sections as below:

1. Opening questions: to introduce the study and the topic, to get to know the participants,

2. Warming up questions: to understand some general concepts and terms relevant to the study,

3. Key questions: to collect the main points of the study and

4. Ending questions: collect additional information, de-brief, and summarise the discussion.

Between 18 October 2021 and 16 December 2021, seven online focus group discussions were conducted with 20 staff members of 15 national CSIRTs with knowledge and experience working in national CSIRTs. To ensure the credibility and reliability of the study design, the focus group schedule was first reviewed with a domain expert from MyCERT, the national CSIRT of Malaysia, to ensure the questions were correct and appropriate (`https://www.mycert.org.my/`). Then, a pilot focus group with four senior staff members of MyCERT was conducted to test the Focus Group Schedule for feasibility, appropriateness and time management [128; 221]. The pilot study allowed for refinement and improvement of the Focus Group Schedule. It should be noted the pilot focus group discussion was not used in the data analysis.

### Recruitment of Participants

The nature of the study required participants to have specific knowledge and a good understanding of national CSIRTs' operations to provide meaningful and rich insights to answer the RQ. Therefore, the study's participants selection was "purposive" instead of "random" [55], consisting of staff members of national CSIRTs. Past studies and experience learned from the research literature showed that the best way to approach potential participants is to rely on the researchers' professional and personal contacts within national CSIRTs and cross-CSIRT organisations [168]. For this study, such contacts were secured via the researcher as an employee of Malaysia's national CSIRT (MyCERT).

A total of 20 participants were recruited through five different channels described below. The researcher recruited four colleagues working at MyCERT as participants. Three participants were recruited using contacts from a previous study reported in Chapter 4 that used the same target community (staff members of national CSIRTs) for participant recruitment, for which they consented to be contacted again for future research. Nine participants were recruited through the researcher's contact with the CERT Division of the Software Engineering Institute (SEI), the Carnegie Mellon University of the USA (`https://www.sei.cmu.edu/about/divisions/cert/`). Three participants were recruited through the researcher's contact with the Organisation of Islamic Countries (OIC) CERT (`https://www.oic-cert.org/`). A final participant was recruited via the researcher's contact with the industry. Official invitation emails were used to recruit participants. A copy of the invitation email is available at `https://cyber.kent.ac.uk/research/CSIRTs/Focus-Group/Email-Invitation.pdf`.

All participants willingly consented to participate in the study, and their participation was not part of their official duties within their national CSIRT role. Four focus groups consisted of three participants each, two more focus groups consisted of 2 participants each, and one focus group had 4 participants. Though four or five participants in each focus group are suggested [141; 89; 130], due to the relatively low number of participants for the study and the difficulties with arranging larger focus groups, the focus groups were kept smaller. Nevertheless, on the positive side, smaller groups meant that all participants had a better opportunity and sufficient time to discuss and elaborate their views during the discussion. Since the study is qualitative, the low number of participants was not considered a significant concern.

How to address researcher-specific bias was considered to ensure the reliability of data collected during the focus group discussion. Notably, the moderator (researcher) is a staff member of a national CSIRT and has professional relationships with four of the 20 participants. To mitigate potential bias caused by such associations, debriefing sessions were held at the end of each focus group discussion to collectively summarise and agree on the points discussed among participants in the focus groups [61]. On the positive side, the study might get better results in this context because CSIRT personnel tend to trust peers from other CSIRTs more, so the focus group discussions could be more open. The list of participants is shown in Table 13. Notably, two participants represented NCSC-NL (the Netherlands), two represented CERT-NZ (New Zealand), and four represented MyCERT (Malaysia). The rest of the national CSIRTs had one participant each.

Table 13: List of National CSIRTs in Each of the Seven Focus Groups Discussion

| **Focus Group (FG)** | **National CSIRT** | **#(Participants)** |
|---|---|---|
| FG 1 | NCSC-NL (Netherlands), Sri Lanka CERT/CC, TunCERT (Tunisia) | 4 |
| FG 2 | BGD eGOV CIRT (Bangladesh), CERT.at (Austria), CERT-SE (Sweden) | 3 |
| FG 3 | MyCERT (Malaysia), CERT-PH (Philippines) | 2 |
| FG 4 | MyCERT (Malaysia), SWITCH-CERT (Switzerland), EG-CERT (Egypt) | 3 |
| FG 5 | MyCERT (Malaysia), CISA-US-CERT (USA), UK-NCSC (UK) | 3 |
| FG 6 | CTIR-Br (Brazil), CERT-NZ (New Zealand) | 3 |
| FG 7 | MyCERT (Malaysia), NCSC-Hungary (Hungary) | 2 |
| | **Total** | **20** |

The 15 national CSIRTs from which participants were recruited can be split into two sub-types. Fourteen have national responsibility for the cyber protection of their respective country or economy. In contrast, one does not have national responsibility for cyber protection of the respective country or economy but specific roles within specific sectors. This is SWITCH-CERT, whose role is to provide incident response services to research and education, the domain registry, and multiple industrial sectors (banks, industry, logistics, and energy) in Switzerland.

### 5.2.2 Data Analysis

The focus group data were empirically analysed using the content analysis method, a qualitative data analysis method explained in Section 1.4.2. Content analysis was adopted in this study as the intention is to capture and quantify new concepts, words and phrases identified in the focus group data to answer RQ3. Content analysis was preferred over a thematic analysis as it allows categorising and describing the data more objectively, aligning with RQ3. Thematic analysis was also ruled out, as there was no intention to identify significant patterns or themes across the focus group data and interpret its underlying meaning.

An inductive coding approach, explained in Section 1.4.2 was adopted to encode the focus group data to highlight interesting concepts, words and phrases more objectively. First, the focus group data was overviewed by counting occurrences of particular attributes for the study and translating them into figures, percentages and histograms. Then, a qualitative approach was adopted to generate the main concepts from the focus group data, from which the significant findings were drawn to answer the study's research question. Notably, for this study, the focus group is the central unit of analysis. Nevertheless, the study also paid attention to individual participants' views during data analysis for inclusion in the findings [22].

Within the content analysis, a bottom-up approach was used to analyse the focus group data to capture participants' opinions, essential for developing systematic procedures or guidelines for evaluating free tools and public data [23]. Similarly, a bottom-up approach was used in creating the McCall Software Evaluation Model [156]. A top-down approach was eliminated as the purpose is to gain as many insights from participants and to engage directly with participants to understand the real-world operations at the national CSIRTs participated in the study [23]. Furthermore, a bottom-up approach is more straightforward compared to a top-down approach.

For content analysis, coding [195] was used to guide the identification of significant information in the data [22; 98] that is of interest to the study. After extracting the codes, they are grouped into several code categories [172] based on commonalities among the codes that align with RQ3.

### Coding of Data

Codes and coding are essential in qualitative data analysis to organise and interpret data and capture emerging concepts, explained in Section 1.4.2. In this study, in-vivo coding was adopted to code the focus group data, as stated in Section 1.4.2. The study adopted a coding model used by Erlingsson et al. (2017) [61] for content analysis. This model was chosen because the steps provided are easy to understand and follow for coding the focus group data. An example of how the focus group data were coded is shown in Table 14.

Before coding, data (transcripts) preparation was done by the following steps:

1. Transcribed the audio recordings of the focus group discussions manually. Transcription software was not used for privacy and accuracy purposes. The researcher listened to the recordings several times to ensure the accuracy of the transcription. Additionally, "member checking" was used to increase the accuracy of the transcription [45].

2. Labelled all transcripts with names of the Focus Groups (FG) – FG1, FG2, FG3, FG4, FG5, FG6 and FG7. For privacy purposes, participants were labelled with the

names of the National CSIRTs they represent. If there is more than one participant from the same national CSIRT, they are labelled as Participant 1 and Participant 2, followed by the national CSIRTs.

3. Loaded all transcripts into a QDA software, Atlas.ti version 8.4.5, one of the most widely used QDA software [147]. The QDA software was used to store transcripts and manage the coding process [100].

4. Data is now ready for the coding process.

After the above steps, transcripts and data are ready for subsequent coding. The coding of data follows the below steps:

1. Transcripts: Read the transcripts in two steps: a) read once for the first time to get an overall understanding and impression of the data, and b) read three times to gain a better understanding of the data and grasp the details and gist of the data.

2. Segments: Highlight and extract the segment of the transcripts that are particularly interesting to the study topic and align with the research question. This is referred to condense the transcripts into segments with significant information relevant to the study.

3. Codes: Generate codes from the segments of the transcripts while keeping the research question in mind [1; 33], as opposed to deductive coding. The lack of related research on the study's topic justifies inductive coding to extract as much information from the data as possible. Additionally, inductive coding is suitable for developing new procedures and guidelines. The type of coding used in this study is In-vivo coding to code the actual spoken words or other specialised words uttered by the participants [22; 76]. This is essential in developing a set of criteria for evaluating tools and data from the focus group data. The codes generated are then scrutinised and refined to ensure consistency.

4. Code categories: A category consists of codes dealing with the same issue or similar meanings. This is often short, factual sounding, with manifest meaning visible in the category and with limited interpretation by the researchers [61]. The codes were grouped using a bottom-up approach by sorting and appraising the codes to determine which codes belonged to a similar group, forming a category aligned with the RQ. Additionally, the codes were also grouped based on their semantic meanings. The categories were then inspected for completeness and redundancy.

Table 14: Coding of the Focus Group Data

| Transcripts | Segments | Code | Group (Category) |
|---|---|---|---|
| Focus Group 4 – P1<br>"We talked to our peers in our vicinity when we do a lot with our European counterparts. We meet about with other teams at FIRST meetings or international conferences and nationally, as well as we talked to fellow teams in the Austria on what they're using, what their experience is and that kind of like gives us the feedback to know, hey, is it worth looking at." | "We talked to fellow teams in the Austria on what they're using and their experience." | Talk to fellow teams | Evaluation Practices |
| Focus Group 1 – P1<br>"Then for tooling, it's all in usability. For tooling, it's important that the needs of the entire goals we have in processing data, in case of sensors and sensor network it's finding problems." | "It's all in usability." | Usability | Criteria Tool Evaluation |
| Focus Group 2 – P1<br>"For tools the main challenge is time if we look through our data, our tracking of what's vulnerable, we get every other day and tool that someone builds somewhere across the world and we just don't have the time to look at all new toys and tools that are popping up globally." | "the main challenge is time" | Time challenge | Evaluation Challenges |

## 5.3 Results

This section presents the results of the study. First, demographic information about the participants will be given, and then the results will be based on the questions in the Focus Group Schedule. It should be noted that the result focuses on qualitative insights from participants' conversations, with minimal quantitative translation of the data into numbers and figures. The study used "P" to refer to participants and "FG" to refer to focus groups. A total of 266 segments were extracted from the seven transcripts of the seven focus groups for subsequent encoding. Then, from the 266 segments, 289 codes that precisely capture the study's topics of interest were identified, keeping in mind the RQ.

The 289 codes were grouped into six code categories based on semantic similarities. The code categories are 1) tool definition, 2) data definition, 3) evaluation practices, 4) tool evaluation criteria, 5) data evaluation criteria and 6) evaluation challenges. Additionally, 66 memos were created during the coding process to facilitate analysis. The code book for the study is available at (`https://cyber.kent.ac.uk/research/CSIRTs/Focus-Group/Codebook.pdf`).

Notably, content analysis method was used to analyse the focus group data and the findings are reflected in Figure 13 and Tables 15–19.

### 5.3.1 Demographic Information About Participants

Participants' roles in their respective national CSIRTs are presented in this section. Most (13) of the participants are involved in day-to-day technical and operational roles. They are the Team Leader, Analyst, Senior Analyst, Specialist, Security Engineer and Executive. Seven participants are involved in managing and administrating the operations. They are the Director, Head, Deputy Head, Advisor and Principal Administrator. The roles of participants are illustrated in Table 15.

Table 15: Participants' Roles Within the National CSIRTs

| Roles | Number of Participants |
|---|---|
| Director | 1 |
| Head | 2 |
| Deputy Head | 1 |
| Team Leader | 2 |
| Principal Administrator | 1 |
| Advisor | 2 |
| Specialist | 3 |
| Senior Analyst | 1 |
| Analyst | 4 |
| Security Engineer | 1 |
| Executive | 2 |

### 5.3.2 Definitions of Tools and Data

The study is interested in understanding how tools and data are defined from the viewpoints of the participating national CSIRTs. The Google Hack Database (GHDB) (`https://www.exploit-db.com/google-hacking-database/`) was used as an example and asked the focus groups to define if GHDB is a data or a tool. GHDB was referred because, in Chapter 4, it was unclear if GHDB is considered a tool or data. Hence, this study clarifies the definition of GHDB. All focus groups (7) agreed and defined GHDB as data consisting of hackers' tactics, techniques, procedures (TTP) and indicators of compromises (IOCs), as well as a database or a data resource to search and collect additional information to understand threats and facilitate incident response.

### Definition of Tools

On the definition of tools, all seven FGs defined tools as follows: "*Tools are essential to support specific functions and purposes in the operations of national CSIRTs, e.g., for network analysis, facilitating incident reporting, processing raw data, bringing value to users, improving existing processes and generating more productive outputs, ultimately achieving users' goals and targets.*"

Notably, one participant pointed out that tools are very loosely defined and interpreted in the literature, and it was timely that this study raised the question. According to another participant, the literature also tends to be imbalanced in defining tools, covering certain types of tools only, e.g., forensic tools, while under-appreciating other tools, e.g., email and ticketing tools.

### Definition of Data

Considering that the research literature supports diversified definitions of data [288], this study is also interested in understanding how national CSIRTs define data.

All seven focus groups defined data as: "*Data can be differentiated based on the access level of the data and how it can be shared with others ( e.g., 'open data', 'public data', 'closed-source data' and 'commercial data'), for communication, to facilitate investigations, obtained from the Internet, trusted partners, vendors, shared and disseminated accordingly.*"

Interestingly, all seven focus groups defined data in the context of incident response. They described data with examples of data such as IP addresses, log lines, botnets, leaked information, and web services. Furthermore, one focus group explained that data must be massaged to identify actionable information.

One participant informed that without data, national CSIRTs would not know what incident to handle, as stated by the participant:

> "*Without this data, we won't know what to do or handle.*" (Interviewee, P1, FG3)

The focus group clarified that a tool needs data as the input to operate, while data needs a tool that converts the data into more meaningful information.

The focus groups also clarified the difference between open data and public data. All seven focus groups defined "open data" as accessible to the general public but have an

Table 16: Detail Evaluation Practices of Tools and Data in Participating National CSIRTs

| Evaluation Practices | Detail Evaluation Practices (Based on FG Data) | FG |
|---|---|---|
| Refer to community reviews | Community and global reviews available on the Internet, interacting with people and peers, getting second opinions from other practitioners, feedback from trusted users within communities | 5 |
| Staff self-check on their own | By trial and error – try number of tools, verifying and checking source codes, trust and confidentiality of data sources, own experiences with the tools and data, results from the tool itself, self-weighing data – sensibility, validity of data, trust and confidence with data and tool, check reputation of data – from trusted people, following law and government regulations on data Privacy & Protection, get impression and rate the data – good or bad | 5 |
| Check with other CSIRTs | Communication with other national CSIRTs – through CSIRT networking, interactions at conferences, meetings and talk to fellow teams within a constituency | 4 |
| In-house evaluation | Conduct proof of concept (POC), through use cases, stress testing, vulnerability testing, functional testing, conduct trial run | 2 |
| Mix of in-house and external party | Have external companies do the testing in addition to in-house evaluation | 1 |

agreement defined by data owners concerning sharing and circulating open data. Meanwhile, “public data” can be accessed, shared, and disseminated freely by the public without restrictions or prior agreements with data owners.

### 5.3.3 Standards and Practices for Evaluation of Tools and Data

The study is interested in identifying the current standards and practices in evaluating free tools and public data in the participating national CSIRTs and the wider community (particularly the relevant industrial sectors) based on participants’ knowledge and experience. The study found that all participants (20) from the seven focus groups were unaware of specific standards or best practices in evaluating free tools and public data. One participant mentioned two standards – ISO/IEC 270001 ‘Information Security Management System” and ISO/IEC:270005 “Information Security Risk Management”. However, as per checking, these two standards are unrelated to evaluating tools or data. One participant also mentioned “Kitemarking” (`https://www.bsigroup.com/en-MY/Kitemark/Kitemark-for-products/`). However, based on checking, “Kitemarking” is frequently used to identify product safety, not to evaluate tools or data.

Additionally, one participant pointed out that there is not much agreement on standards within the cyber security domain compared to other domains, such as the defence sector:

> "*And I guess in terms of standards, I think the problem with a lot of stuff in the threat spaces is that there isn't a huge amount of agreement on standards. It's quite limited, whereas I guess in the defence space when they're talking about how we value by evaluating cyber security or an organisation's cyber security.*" (Interviewee, P3, FG5)

All 20 participants disclosed current practices in evaluating free tools and public data in their respective national CSIRTs, summarised in Figure 13. Most participants (14) mentioned they do not have a systematic, formal and proper process for evaluating free tools and public data in their respective national CSIRTs. A few participants (2) mentioned having an in-house evaluation of tools and data. At the same time, one said the IT Department does the evaluation, and another told of a mixture of internal and external evaluation. This is reflected by the comments below:

> "*Let me confirm this. We also don't have a formal process for evaluating tools or data.*" (Interviewee, P1, FG2)

> "*so we don't really have time for proper, detailed process of tool evaluation.*" (Interviewee, P2, FG7)

> "*For me, I know there aren't any formal procedures like in my opinion, there aren't any, that's obviously why you were doing your research. There are no standards for using software in national CSIRTs.*" (Interviewee, P1, FG1)

This study also revealed how free tools and public data are evaluated in the participating national CSIRTs, summarised in Table 16. Most (5) FGs said they refer to community reviews to identify qualified, free tools and public data. Another five FGs said they self-check the free tools and public data for quality purposes. None of the participants mentioned having a systematic procedure to evaluate free tools and public data in their national CSIRTs.

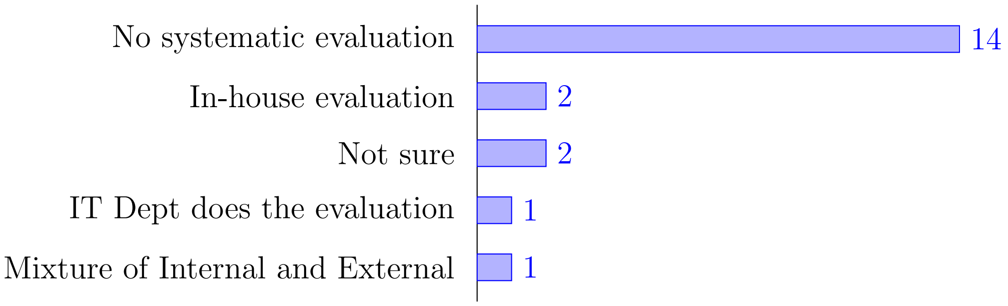


Figure 13: Current Practices in Evaluating Tools and Data in Participating National CSIRTs

Table 17: Challenges and Limitations in Evaluating Tools and Data in Participating National CSIRTs

| Challenges | Detail Challenges (Based on FG Data) | FG |
|---|---|---|
| Lack of awareness & expertise | Lack of knowledge and awareness about tools and data evaluation standards or practices – staff never thought of doing the evaluation, learning new technologies – to facilitate evaluation, dependent on community reviews, knowledge of other evaluation solutions – i.e., riding (piggyback) on the evaluation done by other CSIRTs. | 6 |
| Lack of process & procedures | Do not have formal process – formal procedures, guidelines, standards, no guidance, limited standards for tools and data evaluation, no procedure for OSINT and open-source, done on ad hoc | 5 |
| Lack of Resources | People – Lack of staff and workforce to do the evaluation (staff need to be involved in other tasks, i.e., do Vulnerability Assessment and Penetration Testing (VAPT), involved in investigations, balancing between current workloads and other tasks), operating under small team, need resources for post evaluation (i.e., paperwork – documenting the evaluation, ensure system compliance); budget constraint – to setup evaluation environment (i.e., additional servers, network connectivity) and time constraints – more time is dedicated for incident response, don't have time for proper tools and data evaluation, more time needed to handle a high volume of incident reports. | 4 |

### 5.3.4 Challenges with Free Tools and Public Data Evaluation

The study wanted to understand better the challenges that impinge the evaluation of free tools and public data in the participating national CSIRTs. These challenges are summarised in Table 17.

From the results, most focus groups (6) voiced challenges concerning a lack of knowledge and awareness about tools and data evaluation standards or practices. Furthermore, participants never considered performing tool and data evaluations in the operations. Participants also voiced the need to learn new technologies to facilitate the evaluation of tools and data, which they found challenging. Therefore, participants depend on community reviews and knowledge of other evaluation solutions and even ride on the evaluation done by other national CSIRTs.

Five focus groups revealed their challenge with evaluating tools and data because they did not have a formal process and procedure to guide the evaluations. The focus groups pointed out that currently, there are no specific guidelines, frameworks or standards for evaluating free tools and public data in national CSIRTs, as mentioned by a participant:

> "*There's no framework we will be following.*" (Interviewee, P1, FG3)

And another participant added:

> "*We also don't have a formal process for evaluating tools or data.*" (Interviewee, P2, FG2)

Four focus groups voiced the challenge of lacking staff with small teams running the operations in the national CSIRTs. Besides that, the focus groups commented on the lack of budget and time, impeding free tools and public data evaluation in national CSIRTs. A participant mentions this:

> "*so probably there are some number of tools out there which would be really clever and good and nice to have, but which we just failed to evaluate because of our resource constraints.*" (Interviewee, P2, FG2)

### 5.3.5 Candidate Criteria for Evaluating Free Tools and Public Data

The study is interested in constructing candidate criteria for evaluating free tools and public data from the opinions and views of national CSIRTs. In the McCall Model [156], software quality factors were identified from a literature survey. In contrast, this study identifies candidate criteria based on the opinions of participating national CSIRT staff using several focus group discussions.

Notably, two relevant international standards – ISO/IEC 25010:2011 [104] and ISO/IEC 25012:2008 [103] were reviewed and considered to leverage the candidate criteria identified from the focus group discussions. These standards describe principles for assessing software quality [26] and data quality [269] align with this study. These Standards were adopted in previous studies related to software evaluation [178; 285; 253; 177; 135] and data evaluation [269]. They are considered the best Standards for software quality evaluation compared to other Standards [159]. Additionally, the above-mentioned ISO/IEC Standards are current, not obsolete and were reviewed as valid by the ISO/IEC Standards Committee.

Additionally, this study also reviewed and leveraged other Software Quality Models, e.g., McCall Quality Model [156], Boehm Software Quality Model [28], FURPS Model [81] and Dromey Model [57] apart from ISO/IEC 25010:2011 "Systems and Software Quality Requirements and Evaluation (SQuaRE)". This study leveraged the above Models as these are commonly referenced in the literature [11; 243; 50; 229; 181].

Notably, this study used the term criteria to refer to quality characteristics of a software tool and data, as defined in ISO/IEC 25010:2011 "Systems and Software Quality Requirements and Evaluation (SQuaRE)" [104].

Table 18: Candidate Criteria for the Evaluation of Free Tools

| **Corresponding ISO/IEC 25010:2011 Criteria Term** | **Candidate Criteria (Based on Focus Group Data)** | **FG** |
|---|---|---|
| Usability | Accessibility of tool to CERT, expertise, frees up time, good documentation, good performance tool,how much to train people, interface, limit time of handling incident, solve incident in shortest time, time to produce result, usability, working more efficient result, usability, working more efficient, accurateness of the results, benefit we can get, help our organisation, helps to refine data, make job easier, make our output better, meet our expectation, provide result as close as possible, reputation of tool, tool provide value, trust of tool, trustworthiness, easiest tool, make job more efficient | 7 |
| Maintainability | Updated, well maintained, has a license, active development, has support, can audit failures, can add new capabilities, can fix a tool, active developer, active development, feature request, large community maintaining it | 7 |
| Security | Confidentiality of tool, data we input into tools, privacy handling, integrity of tool, secure feature, free from bugs, source of tool, secure, security feature, secure interaction with all components of a tool, secure development | 6 |
| Functionality | Fundamental capability, fundamental output, functionality of reporting, able to query data, can select type of data, fits for purpose | 3 |
| Compatibility | Has API integration, can be connected to other devices, fit with other tools, suitable for a data format, interoperability | 4 |
| Reliability | Reliable in producing output, availability of tool, accessibility to CSIRT community, can see the source code, continuous funding, Will be with us over the following years, false-positive rates, some less false-positive | 2 |
| Context coverage | Flexible tool, can be customised | 2 |
| Others | Compliance, globally accepted tool, certified by regulatory, product certification, used by the community, used by the security community, certified tool, approval from other countries, approval by legal departments, commonly used by the security community, how large is the community using the tool, how many people are using the tool | 3 |

### Candidate Criteria for the Evaluation of Free Tools

This study identified eight candidate criteria for evaluating free tools from participants' opinions and views, as shown in Table 18. This includes one criterion termed "Others" to the researcher's best knowledge; there isn't the best name (term) and definition that leverages the ISO/IEC 25010:2011 or relevant software evaluation Standards – hence, it is termed "Others". Nevertheless, it would be very interesting if the criterion termed "Others" could be refined into precise and concrete criteria in future work.

The candidate criteria were extracted from the focus group data using the content analysis method following an inductive coding approach. The candidate criteria were then cross-checked against the ISO/IEC 25010:2011 [104] to leverage standardised names (terms) and definitions for the candidate criteria. Candidate criteria identified from this study but not defined in the ISO/IEC 25010:2011 were reviewed and decided for inclusion as candidate criteria. Such candidate criteria contribute to the study's improvement of the current software evaluation models and standards.

All seven FGs mentioned that *Usability* and *Maintainability* are criteria that must be included for evaluating tools. According to the focus groups, "Maintainability" concerns up-to-date, good maintenance and support for tools, fixing issues with the tool, and adding new features. Meanwhile, "Usability" includes learnability, efficiency, trust, effectiveness, accessibility, and user interface.

This is followed by *Security* as mentioned by six FGs, concerns with, for example, confidentiality and integrity of the tool. Four FGs perceived *Compatibility* should be a criterion when evaluating tools, for instance, integration and interoperability. *Functionality* was viewed as a candidate criterion by four FGs, giving examples of having the fundamental capability and fitting the purposes of the tool. *Reliability* was perceived as a criterion by two focus groups followed by *Context Coverage* by two FGs. Notably, *Reliability*, *Flexibility*, and *Automated Report Generation* are consistent with a previous study about elements of good-quality tools [31].

Several candidate criteria identified from the study were grouped as others" as these are not defined in the ISO/IEC 25010:2011. These are *Compliance to security policies*, *Globally Accepted*, *Used by Community*, *Certified by Regulators and Product Certification Services*. Nevertheless, the abovementioned criteria were reviewed and decided for inclusion in the candidate criteria. Among them, "Used by Community" was expressed by participants as an essential criterion for tool evaluation but scarcely mentioned in any standards or models.

### Candidate Criteria for the Evaluation of Public Data

Eight candidate criteria on data quality characteristics from participants' opinions were identified and summarised in Table 19. These candidate criteria are consistent with those

Table 19: Candidate Criteria for the Evaluation of Public Data

| **ISO/IEC 25012:2008 Term** | **Candidate Criteria (Based on FG Data)** | **FG** |
|---|---|---|
| Credibility | Data – legitimate data, reliable data, data integrity, data is secure; source of data – place we download data is trusted, quality of the source of data, reliable source, reputable source, from a trusted partner, useful data | 6 |
| Confidentiality | Less personal identifiable information, data comply to Privacy Act, data is on-premise, policy to protect data in cloud, privacy of data | 4 |
| Currentness | Regularly reported, data is updated | 2 |
| Understandability | Structured data, has data model | 2 |
| Completeness | Backward compatible – There is a history (at least two years) of the same data, contains risk, how important is data, data provides value, data has threat level | 2 |
| Precision | Has information for correlation, data that can be acted, data impacting a region or economy | 2 |
| Accuracy | Less false positives | 1 |
| Efficiency | Machine-readable, Can do right thing | 2 |

defined in the international standard ISO/IEC 25012:2008 “Software product Quality Requirements and Evaluation (SQuaRE) – Data Quality Model” [103]. A similar approach to identifying candidate criteria for evaluating free tools was adopted for identifying candidate criteria for evaluating public data.

Most focus groups (6) perceived *Credibility* as a criterion for evaluating data. Participants viewed reliability, integrity, legitimacy, and data security as important considerations when evaluating data for incident response. Four focus groups of the opinion that *Confidentiality* must be a candidate criterion for evaluating data, which includes data complying with privacy acts and personal information in the data is limited. Other candidate criteria expressed by the focus groups (2 FGs for each criterion) are *Precision* that allows data to be acted upon, *Completeness* related to such as having historical information in the data, *Currentness* related to timeliness of the data and *Understandability* relates to structuredness of the data for easy understanding. *Efficiency* and *Accuracy* (one FG for each criterion) were perceived as essential candidate criteria for evaluating data to ensure data is accurate and can be acted upon. The candidate criteria identified from the study are also consistent with previous studies concerning data quality [269].

## 5.4 Summary

This chapter presented a study using seven focus groups of 20 participants from 15 national CSIRTs worldwide. The study was conducted to understand better the current practices in evaluating free tools and public data in national CSIRTs and to identify candidate criteria for evaluating free tools and public data. The findings from Chapter 4 on how free tools and public data are evaluated are unknown and less understood in the surveyed and interviewed national CSIRTs. This prompted the study reported in this chapter, considering that establishing criteria for evaluating free tools and public data is crucial for national CSIRT operations. Furthermore, all focus groups were positive and receptive to a systematic evaluation procedure for free tools and public data, making this study significant. Content analysis was adopted to analyse the focus group data and derive key findings that answer the RQ.

One of the study's key findings is the candidate criteria for evaluating free tools and public data. However, the candidate criteria's usefulness, deployment and applicability are uncertain and not proven. It is vital to get some certainty that these criteria are useful, ready for deployment and can be applied in practice to evaluate free tools and public data. Hence, more work must be done to ensure the candidate criteria are useful, feasible and applicable. One such work is to validate the candidate criteria for usefulness, deployment and applicability in practice. Therefore, semi-structured interviews were undertaken and reported in the next chapter to validate the candidate criteria with staff members of national CSIRTs for usefulness, deployment and applicability in the operations. Additionally, two sample tools and a data source were evaluated by applying the candidate criteria using several metrics for a more objective validation and how this can help when applying the criteria in actual evaluation practice.

# Chapter 6

# Empirical Validation of Candidate Criteria for Evaluating Free Tools and Public Data Developed for National CSIRTs

This chapter validates the candidate criteria identified in Chapter 5 for usefulness, deployment and applicability in the operational practices of national CSIRTs. The study reported in this chapter was performed to answer RQ4 of this research outlined in Section 1.3. The validation involved two approaches: 1) semi-structured interviews with staff members of national CSIRTs on how they perceive the usefulness and deployment of the candidate criteria in national CSIRTs' operational practices, and 2) a more objective validation of the candidate criteria with two candidate tools and a data source using several metrics for applicability of the criteria in practice. This chapter is structured as follows: section 6.1 introduces the study and original contributions. Section 6.2 explains the methodology adopted for the study, primarily the data collection and analysis. Section 6.3 provides results from the semi-structured interviews and the evaluations of two candidate tools and a data source. Section 6.4 summarises the chapter.

## 6.1 Introduction

Despite the importance of tools and data for national CSIRTs, very few studies exist that discuss free tooling and public data used in their operations. One particular research gap is around evaluating whether these free tools and public data have the necessary quality and meet the usability purposes to support these operations. Shedding light on this research gap can be helpful, especially in ensuring appropriate free tools and public data are selected through systematic evaluation, subsequently enhancing the operations in national CSIRTs [207; 71].

Furthermore, findings from Chapter 4 indicated that free tools and public data are widely used to support incident response in the surveyed and interviewed national CSIRTs. It is known that such tools and data are less often evaluated for quality purposes than commercial tools and data, showing the need to evaluate free tools and public data. This was further investigated and reported in Chapter 5 to gain insights into how free tools and public data are evaluated in national CSIRTs. These insights informed the development of candidate criteria to evaluate free tools and public data for national CSIRTs.

The study reported in this chapter empirically validates the candidate criteria for evaluating free tools and public data constructed in Chapter 5 for practical usefulness, deployment and applicability in practice. Online semi-structured interviews were conducted with nine staff members from nine national CSIRTs from Asia-Pacific, Africa and Europe to gain their opinions on the usefulness and deployment of the candidate criteria in national CSIRTs. After that, two candidate tools and one candidate data source widely used by national CSIRTs were selected, and the candidate criteria were applied to evaluate the candidate tools and data for applicability in practice.

This study revealed three major findings, summarised as follows:

1. All participants perceived the candidate criteria as practically useful for evaluating free tools and public data in the interviewed national CSIRTs.

2. All participants agreed that the candidate criteria could be deployed in the interviewed national CSIRTs and other CSIRTs. Most (eight out of the nine) participants would also recommend the criteria to other national CSIRTs for deployment.

3. The study's attempt to apply the candidate criteria to evaluate two candidate tools and a data source showed that the candidate criteria could be applied in practice to evaluate free tools and public data.

The rest of the study is organised as follows. Section 6.2 explains the methodology used in the study, which includes the data collection and analysis methods. Section 6.3 presents the results of the interviews and the evaluation of two candidate tools and a data source. The results are presented separately. First, the semi-structured interview is followed by the evaluation with candidate tools and data. Finally, Section 6.4 summarises the chapter.

## 6.2 Methodology

The study validates candidate criteria for evaluating free tools and public data constructed in Chapter 5. This is performed by gaining insights into how participants perceive the usefulness and deployment of the candidate criteria via semi-structured interviews. After that, the candidate criteria are applied to evaluate two candidate tools and a data source to examine how the criteria could be converted to more specific metrics. The following subsections explain the candidate criteria, the data collection, the data analysis, and how the candidate criteria were applied to two candidate tools and a data source.

### 6.2.1 The Candidate Criteria

The candidate criteria for this study consist of criteria for evaluating free tools and public data. This study refined the candidate criteria constructed in Chapter 5 into 14 candidate criteria that are more precise and concrete for evaluating free tools. The eight candidate criteria for evaluating public data did not undergo any refinements and remained as from Chapter 5. The candidate criteria for evaluating free tools are divided into two categories – "Product Quality" and "Quality in Use" as per the ISO/IEC 25010:2011 standard. The candidate criteria concentrated on "Product Quality" with their definitions (descriptions) is shown in Table 20. The "Product Quality" category is about *software's static properties and the computer system's dynamic properties.* It covers the following nine criteria: 1) Security, 2) Usability, 3) Maintainability, 4) Compatibility, 5) Functionality, 6) Performance Efficiency, 7) Reliability, 8) Compliance, and 9) Certification. The candidate criteria concentrated on "Quality in Use" with their definitions (descriptions) is shown in Table 21. The "Quality in Use" category is about *the outcome of interaction when a product is used in a particular context.* It covers the following five criteria: 1) Context Coverage, 2) Usability, 3) Effectiveness, 4) Freedom from Risk, and 5) Popularity.

Notably, the candidate criteria for evaluating free tools include three new criteria from this research, which are unavailable in the ISO/IEC 25010:2011 or ISO 9241-11:2018 standards. These criteria are: 1) Compliance, 2) Popularity, and 3) Certification. Compliance and Certification are categorised under "Product Quality" and Popularity under "Quality in Use" since the former is about the static properties of the tools while the latter is about users' experiences after using the tools.

The candidate criteria for evaluating data with their definitions (descriptions) are shown in Table 22. It consists of the following eight criteria: 1) Credibility, 2) Efficiency, 3) Confidentiality, 4) Accuracy, 5) Precision, 6) Understandability, 7) Currentness, and 8) Completeness. These criteria leveraged the ISO/IEC 25012:2008 standard.

It should be noted that most software evaluation models in the literature use the terms "factors" or "characteristics". Nevertheless, this study uses the term "criteria",

Table 20: A Set of Criteria for Evaluating Free Tools (Product Quality) and their Definitions

| Criteria | Definition |
|---|---|
| **Security** | |
| Confidentiality | The degree to which a product or system ensures that data are accessible only to those authorised to have access |
| Integrity | The degree to which a system, product or component prevents unauthorised access to, or modification of, computer programs or data |
| Authenticity | The degree to which the identity of a subject or resource can be proved to be the one claimed |
| **Usability** | |
| Learnability | The degree to which specified users can use a product or system to achieve specified goals of learning to use the product or system with effectiveness, efficiency, freedom from risk and satisfaction in a specified context of use |
| Operability | The degree to which a product or system has attributes that make it easy to operate and control |
| User Interface Aesthetics | This refers to properties of the product or system that increase the pleasure and satisfaction of the user, such as the use of colour and the nature of the graphical design |
| Accessibility | The degree to which a product or system can be used by people with the widest range of characteristics and capabilities to achieve a specified goal in a specified context of use |
| **Maintainability** | |
| Maintainability | The degree of effectiveness and efficiency with which a product or system can be modified by the intended maintainers |
| Supportability | The degree to which a product or system could provide support and assistance to users when encountering problem |
| Analysability | The degree of effectiveness and efficiency with which it is possible to assess the impact on a product or system of an intended change to one or more of its parts, to diagnose a product for deficiencies or causes of failures, or to identify parts to be modified |
| Modifiability | Modifications can include corrections, improvements or adaptation of the software to changes in the environment and requirements and functional specifications |
| **Compatibility** | |
| Interoperability | The degree to which two or more systems, products or components can exchange information and use the information that has been exchanged |
| **Functionality** | The degree to which the set of functions covers all the specified tasks, appropriateness of the tasks and user objectives |
| **Performance Efficiency** | The performance relative to the number of resources used under stated conditions |
| Time behaviour | The degree to which the response and processing times and throughput rates of a product or system, when performing its functions, meet requirements |
| Capacity | The degree to which the maximum limits of a product or system parameter meet requirements |
| Money | The degree of how much money used in relation to the results achieved |
| Human effort | The degree of how much human effort used in relation to the results achieved |
| Material | The degree of how much material used in relation to the results achieved |
| **Reliability** | |
| Reliability | The degree to which a system, product or component performs specified functions under specified conditions for a specified period of time |
| Availability | The degree to which a system, product or component is operational and accessible when required for use |
| **Compliance** | The degree to which tools comply with a specific policy, rules and regulations and operations |
| **Certification** | The degree to which tools are certified and accredited by reputable accreditation and certification bodies |

which reflects the same meaning as “factors” or “characteristics”.

### 6.2.2 Data Collection

Semi-structured interviews [187] were used in the study to draw insights from participants through interactive discussions on how they perceive the candidate criteria. The study received a favourable opinion from the University of Kent’s Central Research Ethics

Table 21: A Set of Criteria for Evaluating Free Tool (Quality in Use) and their Definitions

| Criteria | Definition |
|---|---|
| **Context Coverage** | |
| Flexibility | The degree to which a product or system can be used with effectiveness, efficiency, freedom from risk and satisfaction in contexts beyond those initially specified in the requirements |
| **Usability** | |
| Satisfaction | The degree to which user needs are satisfied when a product or system is used in a specified context of use |
| User experience | The degree of user's perceptions and responses that result from the use and/or anticipated use of a system, product or service |
| Usefulness | The degree to which user needs are satisfied with their perceived achievement of pragmatic goals, including results and consequences of use |
| Trust | The degree to which the user has confidence that the product will behave as intended |
| Comfort | The degree to which user needs are satisfied with physical comfort |
| **Effectiveness** | The degree to which accuracy and completeness with which users achieve specified goals |
| **Freedom from Risk** | |
| Sustainability | The degree to which a system, product or component is sustainable with freedom from risk – economic risk mitigation, health and safety risk mitigation and environmental risk mitigation |
| Harm from use | The degree of negative consequences regarding health, safety, finances or the environment that result from the use of the system |
| **Popularity** | The degree to which the security community (large or small) uses the tool |

Table 22: A set of Criteria for Evaluating Public Data and their Definitions

| Criteria | Definition |
|---|---|
| **Credibility** | The degree to which data has attributes regarded as true and believable by users in a specific context of use. Credibility includes the concept of authenticity (the truthfulness of origins, attributions, commitments) |
| **Efficiency** | The degree to which data has attributes that can be processed and provide the expected levels of performance by using the appropriate amounts and types of resources in a specific context of use |
| **Confidentiality** | The degree to which data has attributes that ensure that it is only accessible and interpretable by authorised users in a specific context of use |
| **Accuracy** | The degree to which data has attributes that correctly represent the true value of the intended attribute of a concept or event in a specific context of use |
| **Precision** | The degree to which data has attributes that are exact or that provide discrimination in a specific context of use |
| **Understandability** | The degree to which data has attributes that enable it to be read and interpreted by users and are expressed in appropriate languages, symbols and units in a specific context of use |
| **Currentness** | The degree to which data has attributes that are of the right age in a specific context of use |
| **Completeness** | The degree to which subject data associated with an entity has values for all expected attributes and related entity instances in a specific context of us |

Advisory Group (CREAG) under the reference number (CREAG071-05-22) on 14 June 2022. All participants willingly gave consent to participate in the study and to have their direct quotes included in research publications resulting from this study, with their personal information anonymised. The Consent Form and the Participant Information Sheet (PIS) used for the semi-structured interviews are available at (`https://cyber.kent.ac.uk/research/CSIRTs/Validation-Criteria/Consent-Form.pdf`) and (`https://cyber.kent.ac.uk/research/CSIRTs/Validation-Criteria/PIS.pdf`),

Table 23: List of National CSIRTs Participated in the Semi-structured Interviews

| National CSIRTs | Website |
|---|---|
| Uganda CERT | `https://www.cert.ug/` |
| Albania CERT | `https://cesk.gov.al/` |
| CERT BUND (Germany) | `https://www.bsi.bund.de/` |
| NCSC Switzerland | `https://www.ncsc.admin.ch/` |
| CERT-MZ (Mozambique) | `https://www.cert.mz/` |
| ID-SIRTII/CC (Indonesia) | `https://idsirtii.or.id/` |
| NCSC-FI (Finland) | `https://www.kyberturvallisuuskeskus.fi/en/our-activities/cert` |
| JpCERT/CC (Japan) | `https://www.jpcert.org/` |
| INCIBE-CERT (Spain) | `https://www.incibe-cert.es/` |

respectively.

The study required specific knowledge, understanding and experiences of national CSIRTs' real-world operations to supplement the necessary information to answer RQ4. Therefore, the selection of participants for the study was purposive instead of random and intentionally selected staff members of national CSIRTs to participate in the study [55]. This was essential to gain accurate, meaningful, and rich insights to answer the research question [55]. Feedback from staff members of CSIRTs is vital to gain insights into the operations of CSIRTs as they have significant experiences in incident responses, much needed when intending to improve CSIRT practices [6].

To recruit participants, six staff members from six national CSIRTs were invited to participate in the study during the 34th Annual FIRST Conference[27] and the [28] in Dublin, Ireland, in July 2022. Three more participants were invited to participate in the study through contacts at the CERT Division of the Software Engineering Institute (SEI) of Carnegie Mellon University in the USA (`https://www.sei.cmu.edu/about/divisions/cert/`). Formal emails were later sent to potential participants to recruit them into this study formally. A sample recruitment email is available at (`https://cyber.kent.ac.uk/research/CSIRTs/Validation-Criteria/Recruitment-Email.pdf`). Finally, nine staff members from nine national CSIRTs willingly consented to participate in the study. A list of the national CSIRTs who participated in this study is shown in Table 23.

The instrument used for this study was an Interview Schedule, which guided the interviewer during the interviews. The interview schedule contains brief information about the interview process, time allocation, and interview questions. It consists of eight open-ended and semi-structured questions, arranged into four sections in sequence: 1) basic information about participants, 2) how staff evaluate tools and data in national CSIRTs, 3) how staff perceive the usefulness and deployment of the candidate criteria

[27] See `https://www.first.org/conference/2022/`

[28] See `https://www.basecybersecurity.com/cyber-security-events-infosec-conferences-it-security-trainings-europe-calendar/natcsirt-2021-2/`

for evaluating tools and data in national CSIRTs, and 4) any other comments about the candidate criteria.

The order of the interview questions was flexible, allowing interviewees to highlight or introduce any other points relevant to the questions. The interview schedule used for the study is available at (`https://cyber.kent.ac.uk/research/CSIRTs/Validation-Criteria/Interview-Schedule.pdf`).

The interviews were conducted virtually via Microsoft Teams (`https://teams.microsoft.com/`) between 2 August 2022 and 6 October 2022. On average, each interview took approximately 30 minutes to complete. With the consent of the participants, the interviews were audio recorded and transcribed. For some interviewees, some follow-up email exchanges took place to clarify their opinions on some criteria. In addition to audio recording the interviews, the interviewer (researcher) also took notes of critical points during the interviews.

To ensure the credibility of data collection, the semi-structured interview questions were reviewed and verified by a domain expert from CyberSecurity Malaysia (`https://www.cybersecurity.my/`), the national cyber security agency of the Malaysian government. A pilot semi-structured interview was held with a senior staff member of the national CSIRT of New Zealand (`https://www.cert.govt.nz/`) on 2 June 2022 following the same setup of the actual nine interviews to ensure the feasibility and appropriateness of the interview questions. The pilot interview also helped improve and refine the interview questions before the interviews. It should be noted that data from the pilot interviews were not used in the data analysis.

### 6.2.3 Data Analysis

Content analysis [276] was adopted to analyse the semi-structured interview data to capture participants' opinions, such as words and phrases, and put them into the study context. Section 1.4.2 provides more information about content analysis. Content analysis was also chosen as it allows for categorising, quantifying and describing the data, objectively. Thematic analysis was ruled out, as there was no intention to explore and identify new themes or patterns across the data and interpret its underlying meaning.

In this study, in-vivo coding was adopted to code the semi-structured interview data, as stated in Section 1.4.2. Erlingsson's coding model [61] was used to guide the coding process. The codes were developed using a data-driven approach (from the raw interview data) instead of theory-driven [160] since the study was not based on any existing theory. This required re-examining the raw data repeatedly to gain more precise insights about the interview data, making the study's code development an iterative process [54]. It should be noted that the researcher coded all of the focus group data. Therefore, there were no issues in establishing consistency or reliability of the coding process (in comparison to a situation in which several coders were involved, which would require further checks to ensure consistency and reliability). Although the study has addressed "between-coder" inconsistency by having one coder, the "within-coder" inconsistency may not have been addressed. To overcome this issue, the researcher reviewed the codes multiple times and iterated before coming up with a final list of codes.

In coding, words, phrases and sentences were extracted from the interview data as meaningful codes while considering the research question [160]. During coding, the focus is on extracting "manifest meaning" (what has been said) instead of "latent meaning" (what is intended to be displayed) [25]. This is because the study needs to capture only surface or explicit meaning from the data. The study was interested in using only words and phrases in the text rather than interpreting the underlying meaning of the words and text [25].

### 6.2.4 Applying the Candidate Criteria

After the candidate criteria were empirically validated using semi-structured interviews for usefulness and deployment, the candidate criteria were validated more objectively using several metrics. This was performed by applying the candidate criteria to evaluate two candidate tools and one data source to derive several concrete metrics and values. Doing so gives further evidence of the practicality of the criteria in practice. This supplements the opinions from the semi-structured interviews and makes the study's findings more credible and reliable.

For each tool and data source, all candidate criteria were checked individually. First, determine if a criterion is relevant to the tool being evaluated. If NO – not relevant, move on to the following criterion. If YES – relevant, identify one or more suitable metrics for the criterion and determine the value for each metric. Value can be derived from 1) factual information about the tool's features from the tool's documentation and 2) output and results obtained after inputting an artefact – a PDF file to the tools.

**Candidate tools and data source.** Two candidate tools were evaluated to demonstrate the discriminatory power of the candidate criteria developed from this research. This was performed by applying the candidate criteria to evaluate two candidate tools – VirusTotal and Hybrid Analysis, using several metrics. Evaluating two candidate tools allows for comparing the tools and highlighting the utility of one tool over the other. Doing so shows that the criteria can be used in practice to compare different tools and guide national CSIRTs to select appropriate tools.

VirusTotal and Hybrid Analysis were selected as the candidate tools for the evaluation exercise due to their significance in assisting national CSIRTs in their daily incident response [260]. This is consistent with an empirical study conducted as part of the thesis presented in Chapter 4, which found that VirusTotal and Hybrid Analysis tools are used in the surveyed national CSIRTs to support incident response. Furthermore, based on the researcher's informal conversations with several national CSIRT staff members, there is indecision in selecting between VirusTotal and Hybrid Analysis tools to best support incident response. Hence, considering the above points, the research decided to evaluate VirusTotal and Hybrid Analysis so the results from the two candidate tools could be discerned. This could guide national CSIRTs in selecting VirusTotal and Hybrid Analysis tools to support incident response.

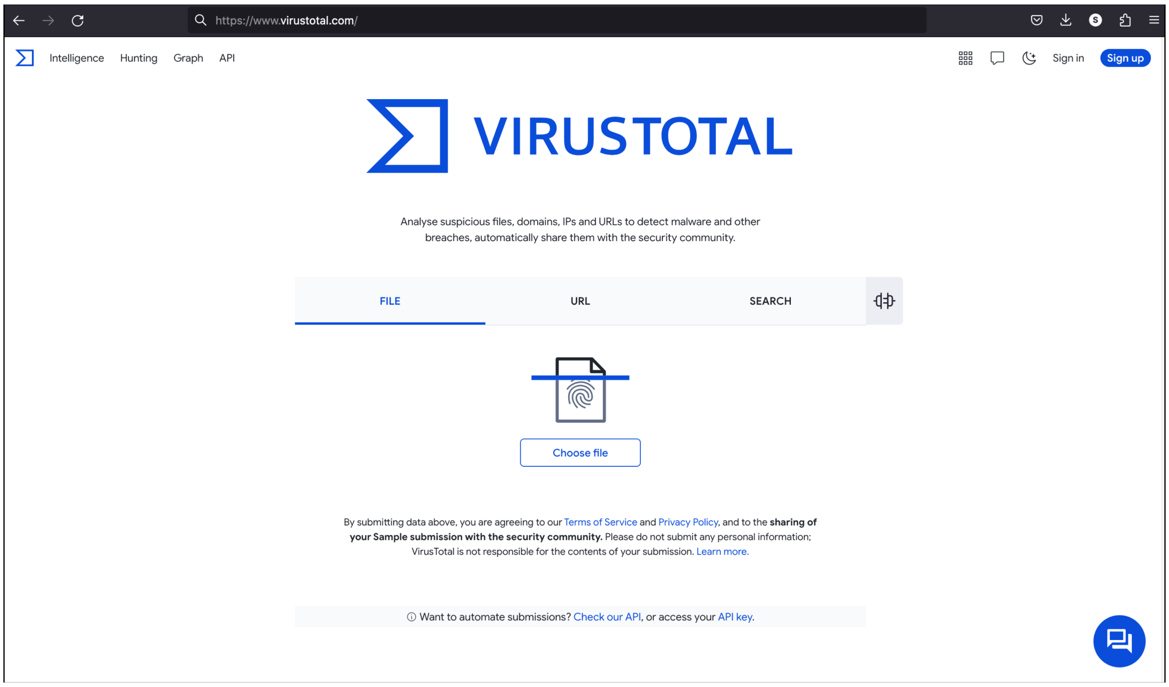
In contrast, only one candidate data source is evaluated due to time constraints. The candidate data source is Shadowserver, which was selected due to its significance to national CSIRTs for incident response work [158; 262; 179]. This is evidenced by its utilisation by some national CSIRTs [118; 133] and also reported in Chapter 3. Furthermore, the researcher accessed the candidate data source through contacts with Malaysia's national CSIRT (MyCERT), facilitating the evaluation exercise.

Notably, the scope of the evaluation exercise for these tools was to submit a sample file to the tools online and observe the outputs. The file can be submitted by clicking the "Choose file" in VirusTotal and "Drag & Drop for Instant Analysis" in Hybrid Analysis and observing the outputs. The Shadowserver data source was evaluated by reviewing the sample data in a CSV file format obtained from MyCERT and observing if the data fulfilled the candidate criteria requirement.

The candidate tools and data source are described briefly below:

1. VirusTotal is a free SAAS (software-as-a-service) tool owned by Chronicle (`https://chronicle.security/`), a subsidiary of Google. VirusTotal can be accessed online to analyse suspicious files, hashes or URLs, which uses several back-end Antivirus engines to facilitate the detection of malware [155]. The tool can be accessed at `https://www.virustotal.com`. It is the largest online anti-malware scanning service, and security researchers widely use it for malware analysis. In a study reported in Chapter 4, VirusTotal is one of the free tools commonly used in the operations of national CSIRTs. National CSIRTs use VirusTotal in daily operations to analyse a file or URL and to confirm if they are indeed malicious. The VirusTotal interface to submit a file or URL is shown in Figure 14.

   VirusTotal inspects files or URLs submitted by users using more than 70 state-of-the-art anti-malware engines and returns engines' detection results if the file or URL is malicious or not [287]. VirusTotal only provides detection results from all the detections employed by back-end engines. As such, the platform does not label the file or URL as malicious or benign; instead, it is up to the user to interpret the detection result provided.



Analyse suspicious files, domains, IPs and URLs to detect malware and other breaches, automatically share them with the security community.

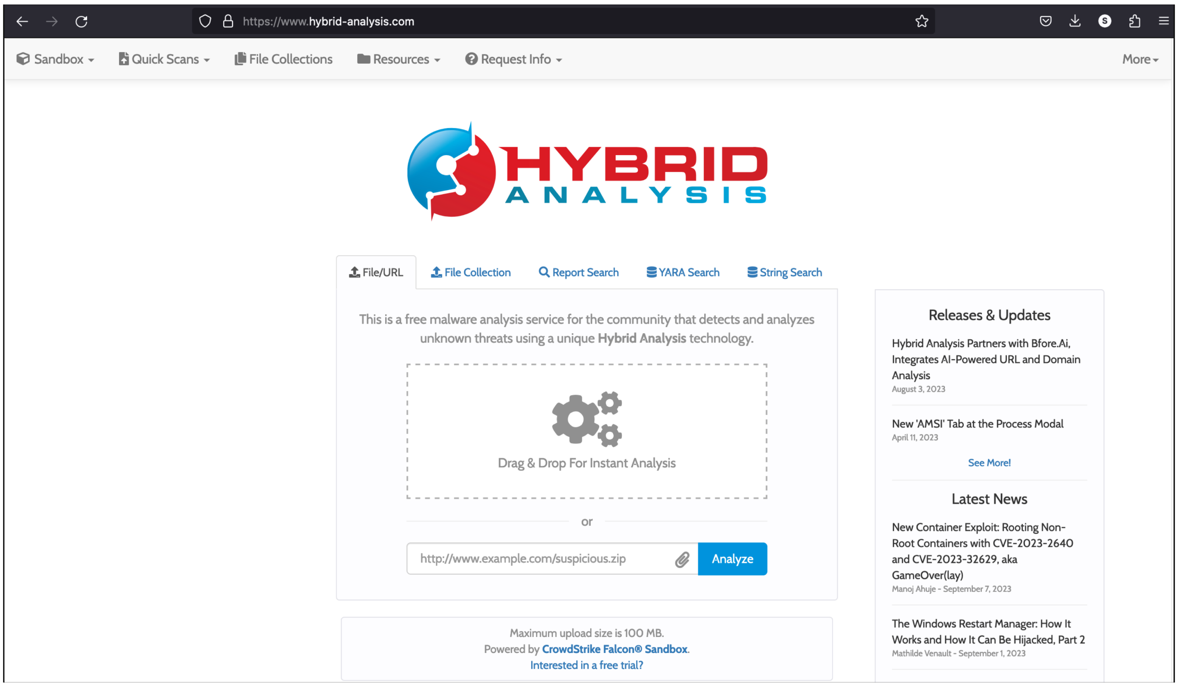
By submitting data above, you are agreeing to our Terms of Service and Privacy Policy, and to the **sharing of your Sample submission with the security community.** Please do not submit any personal information; VirusTotal is not responsible for the contents of your submission. Learn more.

ⓘ Want to automate submissions? Check our API, or access your API key.

Figure 14: VirusTotal Page Interface

2. Hybrid Analysis [56] is a free SAAS tool owned by CrowdStrike (`https://www.crowdstrike.com/`). Hybrid Analysis is a web-based service to detect and analyse malware using a unique Hybrid Analysis technology [211]. The tool can be accessed at `https://www.hybrid-analysis.com`. Hybrid Analysis is an open-source malware analysis platform that can sandbox malicious software and executables. It provides file/URL sandboxing, file collections, reports search, and sandbox results with IOCs and screenshots. In addition, users can search thousands of existing malware reports or download samples and IOCs via the website. It provides instant threat analysis using CrowdStrike Falcon Static Analysis (Machine Learning), reputation lookups, antivirus engines, and static analysis. The Hybrid-Analysis interface to submit a file or URL is shown in Figure 15.

3. Shadowserver (`https://www.shadowserver.org`) data source [227] contains data about malicious Internet activities worldwide (e.g., malware, botnets, spam and computer fraud). The data is structured to include the following fields: date, timestamp, IP address, hostname, geolocation, URL, ASN, port numbers, protocol, name of malware, and file hash. The data is essential for national CSIRTs to notify respective service providers concerning malicious activities originating from their IP addresses or domains. Doing so helps to address emerging threats worldwide and for cyber crime investigations [118]. A sample Shadowserver data source is shown in Figure 16.



https://www.hybrid-analysis.com

Sandbox ▾ Quick Scans ▾ File Collections Resources ▾ Request Info ▾ More ▾

HYBRID ANALYSIS

File/URL File Collection Report Search YARA Search String Search

This is a free malware analysis service for the community that detects and analyzes unknown threats using a unique **Hybrid Analysis** technology.

Drag & Drop For Instant Analysis

or

http://www.example.com/suspicious.zip Analyze

Maximum upload size is 100 MB.
Powered by CrowdStrike Falcon® Sandbox.
Interested in a free trial?

Releases & Updates

Hybrid Analysis Partners with Bfore.Ai, Integrates AI-Powered URL and Domain Analysis
August 3, 2023

New 'AMSI' Tab at the Process Modal
April 11, 2023

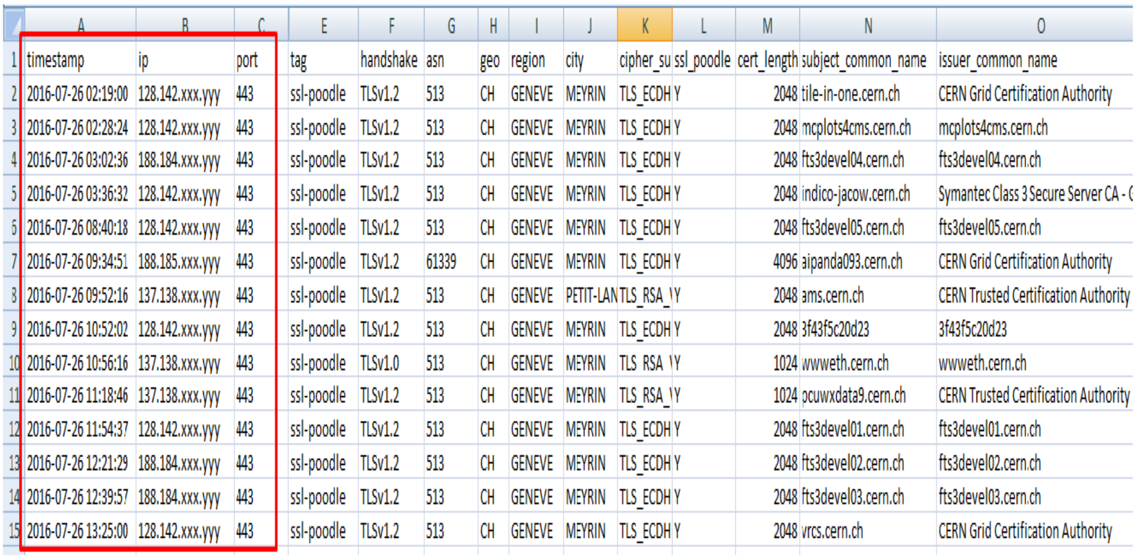
See More!

Latest News

New Container Exploit: Rooting Non-Root Containers with CVE-2023-2640 and CVE-2023-32629, aka GameOver(lay)
Manoj Ahuje - September 7, 2023

The Windows Restart Manager: How It Works and How It Can Be Hijacked, Part 2
Mathilde Venault - September 1, 2023

Figure 15: Hybrid Analysis Page Interface



| | A | B | C | E | F | G | H | I | J | K | L | M | N | O |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | timestamp | ip | port | tag | handshake | asn | geo | region | city | cipher_su | ssl_poodle | cert_length | subject_common_name | issuer_common_name |
| 2 | 2016-07-26 02:19:00 | 128.142.xxx.yyy | 443 | ssl-poodle | TLSv1.2 | 513 | CH | GENEVE | MEYRIN | TLS_ECDH | Y | 2048 | tile-in-one.cern.ch | CERN Grid Certification Authority |
| 3 | 2016-07-26 02:28:24 | 128.142.xxx.yyy | 443 | ssl-poodle | TLSv1.2 | 513 | CH | GENEVE | MEYRIN | TLS_ECDH | Y | 2048 | mcplots4cms.cern.ch | mcplots4cms.cern.ch |
| 4 | 2016-07-26 03:02:36 | 188.184.xxx.yyy | 443 | ssl-poodle | TLSv1.2 | 513 | CH | GENEVE | MEYRIN | TLS_ECDH | Y | 2048 | fts3devel04.cern.ch | fts3devel04.cern.ch |
| 5 | 2016-07-26 03:36:32 | 128.142.xxx.yyy | 443 | ssl-poodle | TLSv1.2 | 513 | CH | GENEVE | MEYRIN | TLS_ECDH | Y | 2048 | indico-jacow.cern.ch | Symantec Class 3 Secure Server CA - ( |
| 6 | 2016-07-26 08:40:18 | 128.142.xxx.yyy | 443 | ssl-poodle | TLSv1.2 | 513 | CH | GENEVE | MEYRIN | TLS_ECDH | Y | 2048 | fts3devel05.cern.ch | fts3devel05.cern.ch |
| 7 | 2016-07-26 09:34:51 | 188.185.xxx.yyy | 443 | ssl-poodle | TLSv1.2 | 61339 | CH | GENEVE | MEYRIN | TLS_ECDH | Y | 4096 | aipanda093.cern.ch | CERN Grid Certification Authority |
| 8 | 2016-07-26 09:52:16 | 137.138.xxx.yyy | 443 | ssl-poodle | TLSv1.2 | 513 | CH | GENEVE | PETIT-LAN | TLS_RSA_\ | Y | 2048 | ams.cern.ch | CERN Trusted Certification Authority |
| 9 | 2016-07-26 10:52:02 | 128.142.xxx.yyy | 443 | ssl-poodle | TLSv1.2 | 513 | CH | GENEVE | MEYRIN | TLS_ECDH | Y | 2048 | 3f43f5c20d23 | 3f43f5c20d23 |
| 10 | 2016-07-26 10:56:16 | 137.138.xxx.yyy | 443 | ssl-poodle | TLSv1.0 | 513 | CH | GENEVE | MEYRIN | TLS_RSA_\ | Y | 1024 | wwweth.cern.ch | wwweth.cern.ch |
| 11 | 2016-07-26 11:18:46 | 137.138.xxx.yyy | 443 | ssl-poodle | TLSv1.2 | 513 | CH | GENEVE | MEYRIN | TLS_RSA_\ | Y | 1024 | pcuwxdata9.cern.ch | CERN Trusted Certification Authority |
| 12 | 2016-07-26 11:54:37 | 128.142.xxx.yyy | 443 | ssl-poodle | TLSv1.2 | 513 | CH | GENEVE | MEYRIN | TLS_ECDH | Y | 2048 | fts3devel01.cern.ch | fts3devel01.cern.ch |
| 13 | 2016-07-26 12:21:29 | 188.184.xxx.yyy | 443 | ssl-poodle | TLSv1.2 | 513 | CH | GENEVE | MEYRIN | TLS_ECDH | Y | 2048 | fts3devel02.cern.ch | fts3devel02.cern.ch |
| 14 | 2016-07-26 12:39:57 | 188.184.xxx.yyy | 443 | ssl-poodle | TLSv1.2 | 513 | CH | GENEVE | MEYRIN | TLS_ECDH | Y | 2048 | fts3devel03.cern.ch | fts3devel03.cern.ch |
| 15 | 2016-07-26 13:25:00 | 128.142.xxx.yyy | 443 | ssl-poodle | TLSv1.2 | 513 | CH | GENEVE | MEYRIN | TLS_ECDH | Y | 2048 | vrcs.cern.ch | CERN Grid Certification Authority |

Figure 16: Excerpts of Shadowserver Data Source [117]

**Sample artefact.** A PDF file was used to input VirusTotal and Hybrid Analysis tools during the evaluation exercise. The PDF filename is *Salinan–Pembayaran–Maybank–Pdf.7z*. The PDF file was obtained through the researcher's employment with Malaysia's national CSIRT (MyCERT). The outputs received after inputting the PDF file to the tools are the values for each criterion evaluated.

The results and findings from the semi-structured interviews and the evaluation exercises with two candidate tools and a data source are presented in the next section. The results from the semi-structured interviews are presented first, followed by the results from the evaluation exercises.

## 6.3 Results

### 6.3.1 Results from Semi-structured Interviews

The results from the semi-structured interviews are presented in three parts: 1) interviewees' demographic information, 2) how participants perceived the usefulness of the candidate criteria, and 3) how participants perceived the deployment of the candidate criteria in national CSIRTs.

Before presenting the participants' demographic information, the coding result of the semi-structured interview data is presented. A total of 66 segments or quotations were extracted from the nine interview transcripts that interest the study. From the 66 segments, 72 codes were derived to answer the RQ. The 72 codes were then merged based on semantic similarity and similarity with the issue the code deals with. After the merging, the codes were arranged into two categories:

- Category 1: Perceptions regarding the usefulness of the candidate criteria.
- Category 2: Perceptions regarding the deployment of the candidate criteria.

The code book for this study is available at (`https://cyber.kent.ac.uk/research/CSIRTs/Validation-Criteria/Codebook.pdf`).

The researcher also created 18 memos reflecting views and thoughts during coding.

#### Demographic Information About Participants

A majority (5) of the participants' roles are managers or heads in their respective national CSIRTs, followed by three as incident responders and one as an executive, as indicated in Figure 17.

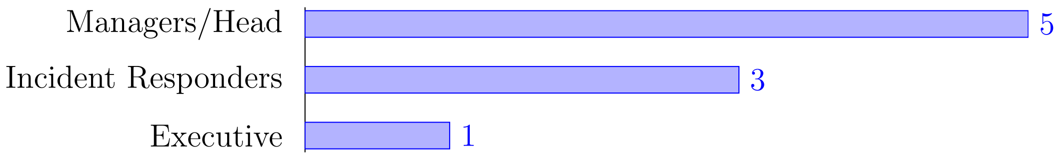


Figure 17: Participants' Roles in Their National CSIRTs

Around half (4) participants have 3- 6 years of working experience in national CSIRTs. Three participants had less than three years, while two had more than six years of experience, as shown in Figure 18.

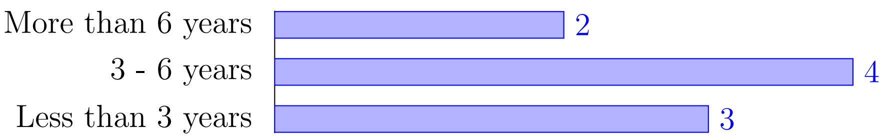


Figure 18: Participants' Work Experience in Their National CSIRTs

### How participants perceived the usefulness of the candidate criteria

All nine participants perceived the candidate criteria as useful for evaluating free tools and public data in national CSIRTs. Most (8) participants perceived the candidate criteria as useful, expressing themselves with the words "good", "nice", and "great". Seven participants said that the candidate criteria could help national CSIRTs select the right tools and data sources, and three mentioned that the candidate criteria were comprehensive and complete. A complete list of opinions captured from participants expressing how they perceived the usefulness of the candidate criteria is shown in Table 24.

One participant found that the criteria used in their current operation for evaluating tools are incomplete compared to the candidate criteria presented in this study. This is because the "quality in use" aspect, reflected from the end user perspective, is missing in their criteria. The participant perceived the criteria from this study could be a good starting point for national CSIRTs looking into evaluating tools and data systematically, as mentioned below:

> "*We have many technical criteria, but not from the end-user perspective. So, I think this is a good, good starting point.*" (Interviewee, NCSC-FI)

Additionally, three participants perceived the approach and method used by this study to develop the candidate criteria as interesting and useful. This is reflected in the comment below:

> "*I guess the general approach is interesting and I think that's quite useful.*" (Interviewee, JpCERT/CC)

### How participants perceived the deployment of the candidate criteria

Furthermore, all (9) participants were willing to adopt the criteria in their national CSIRTs once available as an open resource. At the same time, two participants considered the criteria as a good option for deployment. One participant said they would focus more on deploying the criterion "usability" since it is largely missing in their current criteria. Three participants expressed their willingness to deploy the candidate criteria at any time in their national CSIRTs, as commented by one of them:

> "*Yeah, of course. We are more than willing to use that one.*" (Interviewee, CERT-Albania)

Table 24: Participants' Opinions on the Usefulness of the Candidate of Criteria

| **How Participants Perceived the Criteria for Usefulness (Inductive Codes)** | **Number of Participants** |
|---|---|
| The criteria provided are good, nice, great | 8 |
| Can help national CSIRTs to select tools and data | 7 |
| Useful for operations | 4 |
| Comprehensive and complete criteria | 3 |
| Approach of the criteria is good, helpful and interesting | 3 |
| The criteria are important | 2 |
| Criteria is valuable | 2 |
| A valid research area | 1 |
| Criteria has valid points | 1 |
| Basic idea of the criteria is interesting | 1 |
| The research tackled both sides, tools and data | 1 |
| Methodology used and the evaluation is nice | 1 |
| Needed by the National CSIRTs | 1 |
| The criteria fits | 1 |
| Increase quality of incident response reports | 1 |
| Positive with the criteria | 1 |
| There is no problem with the criteria | 1 |
| Would not take out any points from the criteria | 1 |
| Easy to understand the criteria | 1 |
| Good point | 1 |
| Big help | 1 |

Notably, one participant mentioned that the candidate criteria can be a guideline and best practice for evaluating free tools and public data. The guidelines and best practices may not necessarily be mandatory but recommended for national CSIRTs. This was commented as below:

> "*Saying that the criteria can be a guideline or best practice.*" (Interviewee, CERT-BUND)

One participant expressed their positivity about deploying the candidate criteria in their national CSIRT operation, as it leverages the ISO/IEC standards:

> "*Especially for the criteria, which is already referred to the applicable international standard reports.*" (Interviewee, IDSIRTII)

Three participants perceived the candidate criteria would greatly help, especially the new national CSIRTs when deployed, to select appropriate free tools and public data. This was commented on below by a participant:

> "*Yes, I think I'm sure the tool is very important, especially because we are in our early stages and that can help us use the criteria to actually select the tools that we're going to use.*" (Interviewee, CERT-Mozambique)

Table 25: Participants' Opinions on the Deployment Readiness of the Candidate Criteria for National CSIRTs

| **How Participants Perceived the Criteria for Deployment (Inductive Codes)** | **Number of Participants** |
|---|---|
| Can be used in national CSIRTs | 7 |
| Can be used in CSIRTs | 1 |
| Can be used in all CSIRTs | 1 |
| Worth a try | 1 |
| Can be a best practice | 1 |
| Good if implemented in our organisation | 1 |
| Can be a guideline | 1 |
| To evaluate new tools, of course | 1 |
| It will help us – national CSIRT | 1 |
| We want to borrow the criteria | 1 |
| It is good for new CSIRTs | 1 |
| Beneficial for us | 1 |
| Not a must-have | 1 |
| Don't have criteria like this | 2 |
| Certainly | 1 |

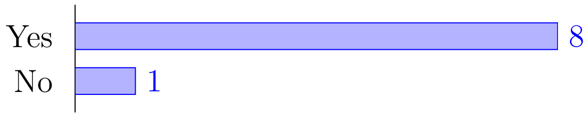


Figure 19: Participants Feedback on Recommending the Candidate Criteria to Other National CSIRTs

This is further supported by another participant who perceived that the candidate criteria would greatly help new national CSIRTs who may not know how to evaluate free tools and public data. This is mentioned below:

> "*Yes, I think this is a very good help, probably for new CERTs who do not know how to evaluate products.*" (Interviewee, NCSC-Switzerland)

A complete list of opinions from participants concerning how they perceived the deployment of the candidate criteria in national CSIRTs is shown in Table 25.

On recommending the candidate criteria to other national CSIRTs, the majority (8) of participants said they would recommend the candidate criteria to other national CSIRTs for deployment, as shown in Figure 19. Only one participant mentioned needing to deploy the candidate criteria in their national CSIRT before recommending it to others. Nevertheless, the participant is optimistic that the candidate criteria would work fine when deployed in their national CSIRT.

### 6.3.2 Evaluation Results of Two Candidate Tools and a Data Source

This section first reports how the candidate tools and a data source are evaluated, followed by the results.

#### Evaluation Result of Two Candidate Tools

The candidate tools evaluated and reported in this chapter are the *VirusTotal* and *Hybrid Analysis*. The tools were manually evaluated between 30 October 2023 and 31 October 2023. All results and observations were recorded in a table format using a Word document on the same day for subsequent analysis.

How the tools are evaluated is described in this paragraph. The criteria for evaluating tools were first reviewed to understand each criterion's requirement(s). Then, the tools' online documentation was reviewed to understand its functions and features. Each criterion was examined to determine if it was relevant to the tools. When a criterion was considered relevant, it was translated into one or more concrete metrics that could cover the requirement(s) of the criterion. The values of the metrics could be 1) Binary (YES/NO), 2) Categorical, 3) Numeric (e.g., the time taken to complete a task), or 4) Descriptive (e.g., the tool supports a GUI).

For metrics that can be quantitatively handled, a sample artefact was fed to the tools to estimate each metric's value. For other qualitative metrics by nature, the value was derived based on the researcher's personal experience and observation of using the tools as a staff member of a national CSIRT – MyCERT. Then, it was determined if the tools fulfilled each criterion's requirement(s) based on the value(s) of the corresponding metric(s).

For example, the criterion "Supportability" is relevant since staff members of national CSIRTs are concerned that they can get support if any issues arise with the tool. The metric used for this criterion is categorical: "what type of support is provided by the tool". After evaluating the tools, the following values were identified: "24x7 Live Support", "Online Support", "Chat Bot" for VirusTotal and "Online Support" for Hybrid Analysis.

A second example is the criterion "Time Behaviour" which is relevant because users are concerned about the speed of obtaining results. The metric definition is: "The amount of time taken to obtain the result when inputting a file or URL (from the start to the return of the results)" and the value is numeric. After evaluating the tools, the following numeric values were identified. For VirusTotal, the value is "2 seconds" and for Hybrid Analysis is "9 seconds".

The evaluation results of VirusTotal and Hybrid Analysis tools are shown in Appendix A. The metrics used for each criterion and the values generated are appended in the above results.

**Evaluation Result of a Candidate Data Source**

The candidate data source evaluated and reported in this chapter is the *Shadowserver* data source. The data source was manually evaluated on 12 February 2023. All results and observations were recorded in a table format using a Word document on the same day for subsequent analysis.

This data source was evaluated by first reviewing each criterion's requirement(s). Then, the data source's online documentation was read to better understand its functionalities. Then, the data source's online documentation was read to understand its features better. Each criterion was examined to determine if it was relevant to the data source, and one or more metrics were identified for each relevant criterion. If the data source fulfils each criterion's requirement(s), it is judged based on the value(s) of the metric(s) identified for the criterion.

One example criterion is "Efficiency". This criterion is relevant for the data source as staff members of national CSIRTs are concerned with the time spent identifying an indicator of compromise (IOC) in the data. The metrics specified for the criterion are numeric: "the amount of time taken to analyse and identify the IOC". The value identified after evaluating the data source is "5 seconds".

A second example is the criterion "Understandability". This criterion was considered relevant for the data source as users are concerned that the data is understandable by staff members. The metric identified is categorical: "the format that human users can understand". The evaluation showed that this metric has two values: "exported as a CSV file" and "displayed in a table". After the evaluation, it indicates the data source can be exported into a CSV file and displayed in a table – for the understandability criterion of the data source.

The Shadowserver data source evaluation result is shown in Appendix B. The metrics used for each criterion and the values generated are appended in the above results.

### 6.3.3 Operationalising the Criteria to Evaluate Tools in National CSIRTs

The evaluation results presented in Appendices A and B show how the criteria should and can be contextualised and translated into concrete metrics to evaluate VirusTotal and Hybrid Analysis tools and Shadowserver data source. This subsection explains how the evaluation exercise with VirusTotal and Hybrid Analysis tools fulfils the requirement to operationalise the criteria in national CSIRTs and demonstrates its discriminatory power presented in the following subsection. This is explained with the following criteria; functionality, usability – learnability, interoperability, supportability, analysability, security, reliability, flexibility, user aesthetics, time behavioural, capacity and popularity. The next section demonstrates the discriminatory power of the criteria to compare different tools – VirusTotal and Hybrid Analysis.

National CSIRTs must ensure tools and data sources have the specified features to meet the national CSIRTs' operational needs and requirements. This requires users to identify a specific feature the tool or data source must possess and perform when required to support the operations. National CSIRTs usually receive incidents regarding suspicious files that might contain malware. Part of the operational task is to scan and check such submitted files if they indeed contain malware. Finding a tool that meets the operational need to scan a file and verify if the file contains malware is essential. Therefore, using the criterion "functionality", national CSIRTs can evaluate if the tool has such a feature. The criterion can be translated into the metric "if the tool has a file scanning feature and performs the specified functionality (feature) accordingly". The value for the metric can be a "Yes" or "No". Hence, national CSIRTs can use the criterion and metric to evaluate if the tools indeed possess the specified feature to scan a file, are workable and provide results that meet national CSIRTs' objective.

It is also essential for national CSIRTs to ensure the usability of tools to support the operations. Two aspects are exemplified to indicate how the criterion "usability" can be operationalised into concrete metrics. One aspect of usability is about learning how to use a tool to conduct a specific task. The criterion "learnability" can be used to evaluate how easily staff members can learn the task of submitting a suspicious file to the tool, view and understand the results – i.e., which button to click for submitting the file, what format of file can be submitted, the file size allowed for submission, where and how to view the result. This criterion can be translated into the metric "amount of time a typical end user (such as a member of technical staff at a national CSIRT) takes to learn the task of submitting a file to the tool in order to get the result or output after submitting the file". The value for this metric is the "time taken which can be in seconds, minutes, hours or days". This can help national CSIRTs choose a tool that meets their objectives in terms of how much time it takes for staff members to learn how to use a tool to conduct a specific task.

Another aspect of usability is the attractiveness of tools. This focus on simplicity which is essential to allow staff members to navigate the tool and view the result in a simplified manner. Hence, the criterion "user aesthetics" can be used by national CSIRTs to evaluate the attractiveness and simplicity of tools. The criterion can be translated into metrics such as "how the tool's result is presented". The values for the metric can be whether the "result is presented in a table or in a different format", "if there are colour codes that can differentiate different aspects of the result", or "if there are icons or buttons to view more detailed results". This is important for staff members' experience with the tool which is simplified for easier navigation of the tool.

Another critical aspect of tools that needs consideration is that they are compatible and interoperable. This is by assessing if the tool can be integrated with other tools by allowing free data flows between them. National CSIRTs can use the “interoperability” criterion to evaluate the interoperability of tools. The criterion can be translated into a metric such as “if the tool allows integration and information exchange with third-party applications within the national CSIRT or workflows”. The value for the metric can be “if the tool has an API for integration with third-party applications” or “has an export feature that allows data exchange”. This ensures the tools used in the operation are interoperable with other existing third-party tools. Doing so helps avoid a mismatch with existing hardware, operating systems, networks or technologies in national CSIRTs.

Notably, technical support of a tool is essential for the operations of national CSIRTs. This is to ensure the tool provides appropriate support to troubleshoot and fix any issues that might arise with the tool. The criterion “supportability” can be used by national CSIRTs to assess the tool’s support mechanisms. The criterion can be translated into a metric such as “what type of support is provided by the tool”. The values for this criterion can be “24x7 live support”, “online support” or even “online chatbots”. Having tool support or even multiple supports shows that the tool consistently provides the assistance required to resolve issues and continuously improve the utilisation of the tool.

Notably, it is also essential that the causes of failures encountered while using the tools can be analysed. The criterion “analysability” can be used by national CSIRTs to evaluate if causes of failures can be identified. The criterion can be translated into a metric such as “if causes of failures can be identified”. The value for this metric could be a “Yes” or “No”. This can be assessed from the tools’ documentation or FAQs that provide information that could inform potential reasons for the failure. This allows for a re-do of the evaluation based on the guidelines provided in the documentation or FAQ to avoid possible failures and achieve the desired output from the tools.

In the operation of national CSIRTs, when choosing tools, it is also essential to consider if a tool complies with security. One aspect of security is data integrity (such as data is not sniffed, modified or manipulated by a third party). This ensures the integrity of data is protected. This aligns with the requirement of national CSIRTs to ensure the data they deal with has integrity. The criterion “integrity” can be used by national CSIRTs to evaluate if the tools comply with security requirements. The criterion can be translated into a metric such as “if the data transmission is encrypted end-to-end using HTTP”. The value for this metric could be a “Yes” or “No”. This ensures the tools used in national CSIRTs meet the security requirements and standards.

National CSIRTs need to ensure a tool is flexible (without modifying the original code of the tool) to the needs of specific users, such as non-expert users, to achieve goals and targets. The criterion “flexibility” can be used by national CSIRTs to evaluate if staff members perceive the tool as flexible for non-expert users. The criterion can be translated into a metric such as “if users perceive the tool is flexible for non-expert users”. The value for this metric could be a “Yes” or “No”. These perceptions can be obtained from estimations with non-expert co-workers from a national CSIRT.

It is essential that a tool is always available to users to enable them to conduct day-to-day operations within the national CSIRT. The criterion “availability” can be used by national CSIRTs to evaluate if a tool is available to perform the tasks whenever required. The criterion can be translated into a metric such as “if the tool can run if any third-party libraries or data sources it relies upon are down”. The value for this metric can be a “Yes” or “No”. This ensures the operations of national CSIRTs are not interrupted because of the unavailability of tools to perform their tasks due to dependencies on other applications or add-ons. Another metric that national CSIRTs can use to evaluate the availability of tools is “to test the uptime of the tool to perform its intended function without failure”. The value for this metric could be “99.999 percentage of uptime”.

The criterion “reliability” can be operationalised by national CSIRTs to evaluate the accuracy of the tool’s detection. The criterion can be translated into a metric such as “the tool returns accurate results for the given input” by testing a set of malware with golden vectors to see what is the detection accuracy of the tool. The golden vectors are essentially benign files that are used as a baseline for comparison during malware testing. For example, evaluating a virus scanning tool for “reliability” should return accurate results that indicate a malicious file is indeed detected malicious after inputting the file. The value for this metric could be a “Yes” or “No”. At the same time, testing by inputting a benign file should return an accurate result indicating the benign file is clean.

Besides their reliability, it is also vital that tools used to support incident response meet specific users who can access the tool for investigation purposes. The criterion “accessibility” can be used by national CSIRTs to evaluate the accessibility of tools. The criterion can be translated into a metric such as “if access to the tool is enabled for users with cognitive or physical impairments”. The value for the metric can be that “the tool’s sound control and brightness can be adjusted” or “the font size of the web page can be enlarged” to suit the users’ needs. This ensures the tool can be accessed even by staff members with diverse abilities and physical impairments, to meet the objectives of national CSIRTs for incident responses.

An up-to-date and continuously maintained tool is crucial to support and meet the operation’s needs and requirements. The criterion “maintainability” can be used by national CSIRTs to evaluate if a tool is maintained. The criterion can be translated by a metric such as “if the tool shows the current year of the tool, visible to users”. This is by checking the current year shown on the tool’s webpage. The value could be showing the “current year 2024”. A well-maintained tool can provide valuable insights into the evolution of the tool and the level of up-to-date based on the year.

The criterion “effectiveness” can be used by national CSIRTs to evaluate if the tool provides a complete and accurate result that fulfils their target. The criterion can be translated into a metric such as “if the tool fulfils the specified requirements and performs the specified functions correctly”. For example, if the tool’s task is to scan a malicious file and returns a result showing that the file contains malware”. The value for this could be a “Yes” or “No”. This is important to ensure tools used in the operations are effective and perform the functions as intended.

Apart from effectiveness, it is also essential for national CSIRTs to ensure that tool providers sustain the tools without depletion (with the fundamental feature). The criterion “sustainability” can be used by national CSIRTs to evaluate the sustainability of the tool. The criterion can be translated into a metric such as “how the sustainability of the tool is guaranteed”. The value for this metric could be that “a large, well-known organisation maintains the tool”, “dependent on a single company’s existence and business-driven decision”. Evaluating a tool for sustainability is vital to allow continuous use of the tool in the operation.

It is also critical to ensure the efficiency of tools that will be used to support incident response in national CSIRTs. The criterion “time behaviour” can be used to evaluate how fast staff members can get the desired result with a tool after inputting a file. The criterion can be translated into a metric such as “the amount of time taken to obtain the result when inputting a file (from the start to the return of the results”. The value for the metric is the time, which can be in “seconds, minutes, hours or days”. This is essential for national CSIRTs to ensure tools can provide quick results, yet reliable for incident response purposes in the operations.

A tool’s capacity or scalability is also essential when selecting tools that meet the operation’s needs and requirements. The criterion “capacity” can be used to evaluate the file size a tool can accommodate for file inputs. This is essential to ensure the tools used in the operations are scalable and can accommodate large file sizes. Sometimes files that need to be investigated in the operations are large. The criterion can be translated into a metric such as “maximum file size that can be uploaded for analysis’. The value for this metric is the “file size, typically in kilobytes (KB) or Megabytes (MB)”. The values can be obtained from the tools’ documentation.

A tool’s popularity is also essential when selecting tools that meet the operation’s needs and requirements. Generally, staff members would like to know if a tool is widespread and largely used in other national CSIRTs or security organisations. The criterion “popularity” can be used to evaluate the popularity of tools. The criterion can be translated into a metric such as “how many security organisations or national CSIRTs use the tool”. The value for this metric can be ‘Large’, ‘Medium’, or ‘Low’. This value can be obtained from external reports and informal discussions with the security community.

Overall, the evaluation exercise showed how the criteria can be contextualised into several metrics and operationalised to evaluate tools. Operationalising the evaluation criteria is essential so national CSIRTs can determine the best tools for incident response after the evaluation exercise.

### 6.3.4 Demonstrating the Criteria's Discriminatory Power to Compare Different Tools

This subsection explains how the criteria – learnability, time behaviour, capacity and popularity, can be translated into quantitative metrics to allow a simple comparison of different tools and demonstrates the criteria's discriminatory power. Quantitative metrics provide objective data that can be used to evaluate and compare different tools. The two tools that were compared during the evaluation exercise are VirusTotal and Hybrid Analysis.

When choosing a tool, it is vital to consider its learnability. For example, how easy staff can learn to use a tool, and if it saves time in learning to use a tool. The criterion "learnability" can be translated into the metric of how much time it takes to learn the task of submitting a file to the tool and obtaining a result. In this exercise, different values were obtained concerning the time taken to learn a task. It took 5 minutes to learn how to submit a file to VirusTotal and get a result, while it took 8 minutes to learn how to submit a file to Hybrid Analysis and obtain a result. Using the criterion, the tools can be compared in terms of how much time it took to learn the task of submitting a file to the tools and obtaining the result. This indicates by translating the criterion into specific metrics, the tools can be differentiated on the learnability aspect. Hence, National CSIRTs can decide which tool has the best learnability option that meets its objective after comparing them using the criterion.

It is also critical to ensure the efficiency of tools that will be used to support incident response in national CSIRTs. Therefore, a tool's time behaviour and capacity must be considered when selecting the tools. The criterion "time behaviour" can be used to evaluate how fast the tool can generate the desired result. The evaluation exercise generated different values when evaluating with VirusTotal and Hybrid Analysis. VirusTotal generated a value of "2 seconds", while Hybrid Analysis generated a value of "9 seconds". Using the criterion, the tools can be differentiated from the values generated for the concerned metric. This allows national CSIRTs to select a tool that offers time behavioural options that meet their objective after comparing the tools using the criterion "time behaviour".

The criterion "capacity" can be used to evaluate the file size each tool can accommodate. The criterion can be translated into a metric such as "maximum file size that can be uploaded for analysis'. The evaluation generated different values when evaluating the maximum file size VirusTotal and Hybrid Analysis can accommodate. VirusTotal generated a value of "650 MB", while Hybrid Analysis generated a value of "100 MB". Using the criterion, the tools can be differentiated by the values generated for the maximum file size each tool can accommodate. After comparing the tools using the criterion, it provides an option for national CSIRTs to select a tool that accommodates their needs.

A tool's popularity is also essential when selecting tools to meet the operation's needs and requirements. Generally, staff members would like to know if a tool is widespread and largely used in other national CSIRTs or security organisations. The evaluation generated different values when evaluating the popularity of the tools using the metric "how many security organisations or national CSIRTs use the tool". VirusTotal generated a value of "large", while Hybrid Analysis generated a value of "medium" regarding the number of security organisations or national CSIRTs using the tools. Using the criterion, both tools can be differentiated in terms of popularity from the values generated. After comparing the tools using the criterion, it provides an option for national CSIRTs to select a tool that meets the objectives of the national CSIRT.

The evaluation of VirusTotal and Hybrid Analysis showed how the criteria could be operationalised to compare and distinguish two different tools. Hence, the usability of one tool over the other can be discerned. The evaluation could enable informed decisions on selecting suitable tools aligning with the operation's objective. Furthermore, it could help compare and identify a tool's strengths, advantages, and prospects, which can help achieve the overall operations goals.

Concerning criteria for evaluating data sources, a similar analysis can be done to demonstrate its discriminatory power, which can be potentially undertaken for future work.

## 6.4 Summary

This chapter presents an empirical validation of candidate criteria developed in Chapter 5. The empirical validation gives certainty on the validity of the criteria before it can be deployed for evaluating free tools and public data in national CSIRTs' operations. Nine semi-structured interviews with nine national CSIRTs staff members were undertaken to get insights into how staff members perceived the usefulness and deployment of the candidate criteria in national CSIRTs' operations. The result from the semi-structured interviews found all participants perceived the candidate criteria as useful and ready for deployment in national CSIRTs. The candidate criteria were further objectively validated by evaluating two candidate tools and one data source using several metrics. Doing so ensures the study's rigour. The result from the evaluation exercise showed the criteria are practically applicable for evaluating free tools and public data in practice.

This eventually addresses the gap reported in Chapter 4 and Chapter 5 – the lack of a systematic procedure for evaluating free tools and public data in national CSIRTs' operational practices. The validated criteria will be released as an open resource and recommended to national CSIRTs for real-world application to evaluate free tools and public data. Cross-CSIRT channels and platforms like FIRST, ITU, and ENISA could promote the criteria to national CSIRTs and the wider CSIRT community for practical application in evaluating free tools and public data. The evaluation results of the candidate tools and data source will be released as an open resource that could serve as a case study. This allows national CSIRTs to have more concrete examples of evaluation exercises and guide the deployment of the candidate criteria in practice.

# Chapter 7

# Discussion

This chapter first discusses the research findings presented in Chapters 3, 4, 5 and 6. Key insights regarding using free tools and public data in the operational practices of national CSIRTs are discussed in Section 7.1. The limitations of this research and measures to overcome these in ways that do not significantly impact the overall findings are discussed in Section 8.3. Finally, Section 7.2 summarises the chapter.

## 7.1 Key Findings

### 7.1.1 State of the art regarding the use of free tools and public data in the operations of national CSIRTs – Towards RQ1

The key findings from Chapter 3 are discussed below. These are findings from an SLR that was conducted to gain insights into the state-of-the-art regarding the use of free tools and public data in the operational practices of national CSIRTs. The findings are towards answering RQ1.

The general information from the SLR shows this topic is researched and published almost yearly and worldwide but in a small volume. Though in small volume, it suggests the topic of national CSIRT operational practices is relevant to researchers and incident response communities worldwide – indicating a need to further progress this research area. Some areas have already been identified, including the fact that literature around the use of free tools and public data in national CSIRTs is still generic, incomplete, ad hoc, or fragmented – indicating a research gap. On the other hand, more helpful information about adopting free tools and public data in the operations of national CSIRTs can be gathered from the websites of national CSIRTs and cross-CSIRT organisations compared to research papers – further evidence of the research gap identified above.

Notably, in many cases, publicly available information on national CSIRTs that is derived from their websites is incomplete and not meant to disclose which tools are used in their operations. In some cases, there may be some indirect indication that national CSIRTs use the mentioned tools, but no explicit representation because this is not the purpose of the documents published on their websites. This justifies why the list of tools and data identified from this research may be incomplete, with concrete findings calling for more empirical studies on this topic.

From the websites of national CSIRTs and cross-CSIRT organisations, the mixed-use of public and closed-source data is common in national CSIRTs. Even commercial and free tools consume such data for cyber threat investigation purposes. Similarly, the mixed-use of free and closed-source tools is also common in national CSIRTs, and many free tools are recommended for daily operations. The usage is primarily concentrated in several areas, such as information sharing across national CSIRTs, general cyber threat intelligence, digital forensics and malware analysis. Several public data sources and security feeds play essential roles in the operations, including CVE, NVD, malware information, attack sensors, darknet, and OSN such as X. This implies that national CSIRTs use public data sources for incident response. However, an open discussion on how it is used and perceived is unknown from the literature.

Public data and free tools are ubiquitous within national CSIRTs, probably because of the free and open-source nature of much relevant software such as Linux distributions and digital forensics tools. This can be seen from significant software projects: a) MISP with ongoing development by CIRCL, b) TARANIS developed by NCSC-NL and c) IntelMQ developed by CERT.at. This complements the encouragement for national CSIRTs to seek tools developed by the CSIRT community to prevent duplication and to standardise communication among national CSIRTs [111]. Several national CSIRTs (e.g., CIRCL, JPCERT/CC) also provide public-facing open data and free tools to other national CSIRTs. This aligns with recommendations for national CSIRTs to prioritise open technology as it is free, independent and cost-saving [88]. Hence, the continuous usage and advocating of free tools and public data in national CSIRTs indirectly give the impression that staff members of national CSIRTs are optimistic towards free tools and public data. Many national CSIRTs recommend and advise organisations and citizens on end user-facing free tools, indicating their usefulness in supporting incident response operations without dependence on commercial tools.

### 7.1.2 Empirical evidence that free tools and public data are used in national CSIRTs operations – Towards RQ2

The key findings from Chapter 4 are discussed below. These are findings from an empirical study conducted to gain insights into the real-world operational practices of national CSIRTs regarding the use of free tools and public data. The findings are towards answering RQ2.

OSINT, free tools, and public data are common among most (if not all) national CSIRTs who participated in the study, independent of attributes such as region, size, economic situation, and maturity level. Although the survey and the interviews did not provide a complete set of OSINT, free tools, and public data, the participants' responses imply that a wide range is used within the participating national CSIRTs. Notably, it was found that the decision to select and use specific tooling was generally informal and ad-hoc. Furthermore, less is understood about how these toolsets can best be evaluated to ensure they are effective for various operations. This presents an operational gap in current practice and indicates the need for further research into how free tools and public data are evaluated in national CSIRTs.

There is also general agreement among all participating national CSIRTs that closed-source data are insufficient to support their incident investigation within national CSIRTs. This means there is an urgent need for good-quality public data to supplement closed-source data for more effective incident response. Most participants perceived that public data are valuable as additional sources to provide richer information and intelligence, especially when combined with closed-source data, which is commonly insufficient. The research also found that most national CSIRTs who participated in the study faced challenges in obtaining details of reported incidents, which means there is a lack of closed-source data to facilitate analysis in national CSIRTs. This is due to a lack of reporting and sharing of details by victims and organisations experiencing incidents to national CSIRTs. This suggests why national CSIRTs often consider closed-source data insufficient or incomplete for further incident investigation. This observation also provides a partial answer as to why it would be necessary for national CSIRT staff to acquire more information from public data to fill the insufficiency in closed-source data.

Public data's need and proven usefulness justify the wide use of OSINT tools in the participating national CSIRTs. Significantly, many tools used in the participating national CSIRTs are free and open-source for both OSINT and non-OSINT purposes. Interestingly, one participant pointed out that some free tools are comparable to commercial tools for similar functions. This gives assurance that it is possible to operate a national CSIRT without dependencies on (expensive) commercial tools. This finding inspires start-up national CSIRTs and those national CSIRTs running with budget constraints.

In addition to the positive aspects of public data, OSINT, and free tools discussed above, participants also commented on the operational challenges of using OSINT, free tools, and public data. Among all the challenges, two are particularly significant in informing future studies on this topic. First, national CSIRTs face challenges in terms of the validity and usability of tools for incident responses. The second challenge is the reliability and understandability of public data, which is often found unstructured and therefore does not easily lend itself to conducting analysis. Future work is suggested to address the above two challenges. Such work shall inform the development of systematic procedures or guidelines for evaluating free tools and public data for national CSIRTs and, eventually, identifying qualified and usable tools and data to support incident responses more systematically and subsequently, improving the current practice in national CSIRTs and broader security operations to a more systematic approach.

### 7.1.3 A Set of criteria that can be used to evaluate the suitability of free tools and public data in incident response practices of national CSIRTs – Towards RQ3

The key findings from Chapter 5 are discussed below. These are findings from a focus group study about how free tools and public data are evaluated in national CSIRTs and a set of candidate criteria for evaluating free tools and public data. The findings are towards answering RQ3.

To begin with, the research intended to understand better how staff members of national CSIRTs define tools and data from their knowledge and experience, besides how the literature defines them. It is clear from the study that the tools and data expressed pragmatically by the study participants are based on their knowledge and experience as staff members of national CSIRTs. This gives the impression that the participants have a sound understanding and knowledge of the tools and data they use daily to facilitate incident responses. The study also found that tools and data depend on and complement each other, playing vital roles in the operations of national CSIRTs. Notably, the definition of data derived from the study is consistent with how the literature defines data, though, in principle, the definitions can be pretty diverse [288]. Significantly, all national CSIRTs who participated in the study had a shared understanding of what constitutes tools and data without contradicting one another. Such a common understanding is crucial for timely threat information sharing, tools and data exchanges among national CSIRTs, subsequently helping to improve overall incident responses. It is also worth mentioning that there is no consensus among all focus groups on what is defined as tools and data. Instead, these are described in fragments but consistent across all focus groups and participating national CSIRTs.

Another noteworthy finding is the lack of knowledge among the participating national CSIRTs about standards and best practices for evaluating tools and data. This implies that 1) the ad hoc practices of tools and data evaluations in the participating national CSIRTs and 2) a lack of systematic procedures for evaluating tools and data in the participating national CSIRTs, in agreement with findings reported in Chapter 4. Nevertheless, it is enlightening that participants are optimistic about having more systematic procedures for adopting free tools and public data in national CSIRTs. This is proven by the focus groups' eagerness to know the study's outcome, implying that national CSIRTs are receptive to systematic procedures for improving current practices. These are reflected in some candid comments by participants.

The findings are also consistent with a recent study [27], which stressed the need to evaluate tools, particularly open-source tools, to ensure the sustainability and success of national CSIRTs' operations. This calls for more research to establish systematic procedures to evaluate tools and data for national CSIRTs, benefiting the broader security community. Surprisingly, it was revealed from the study that other national CSIRTs also desire to develop an approach for evaluating tools and data by getting collective opinions from other national CSIRTs. A participant told this:

> "*When it comes to evaluating, I'm not sure how that is organised in NCSC-XX, but I know that we are still desiring a more international approach to evaluating the tools that we use in the community and also to get ideas from other countries and other national CSIRTs and to share our best practices.*"
> (Interviewee, P2, FG1)

This indicates that the research presented in this thesis is particularly interesting to national CSIRTs and is vital for the national CSIRTs community. Furthermore, the approach used in this research is getting collective opinions from national CSIRTs to build the procedures envisaged by national CSIRTs. Hence, it is timely for the research to be undertaken and fulfil the expectations of national CSIRTs.

There is also a general agreement among all focus groups about the challenges in tools and data evaluation in national CSIRTs. Among all the challenges, one is particularly important: *lack of formal process and procedures for evaluating tools and data* – indicating an operational gap. The challenge highlighted is essential as it gives the impression that establishing a systematic approach for evaluating tools and data is beyond national CSIRTs' capability – due to a lack of expertise, staffing, resources and budget. This shows why independent research is necessary to improve current practices in national CSIRTs. One such study is conducted and presented in Chapter 5. Looking into the broader picture, this also indicates an urgent need for stakeholders and researchers to develop more cyber security-related procedures.

The constructive opinions and interactive discussions among the focus groups on deriving candidate criteria for evaluating free tools and public data give the impression that national CSIRTs collaborate with researchers to enhance operational practices. Such positivity is vital to advance this research with open minds through interactive discussions among the focus groups to develop candidate criteria for evaluating free tools and public data in the operations. Subsequently, it helps national CSIRTs to make sound evaluations of free tools and public data compared to existing practices. It also gives the impression that such interactions and collaborations among national CSIRTs could be undertaken in the future to enhance national CSIRTs' operations. It is also worth mentioning the candidate criteria that emerged based on participants' practical knowledge and experiences on what should constitute quality tools and data- pragmatic rather than theoretical. This implies considering staff members' opinions of free tools and public data, vital to enhancing national CSIRTs' operations. This also gives the impression there is a very high likelihood that the criteria can be applied in practice, but this needs to be proven. It might be that the candidate criteria from this research are not a complete set. Still, they provide a solid foundation for further refinements and extensions in future work for a more extensive and inclusive set of criteria.

### 7.1.4 Perception on the usefulness, deployment and applicability of the proposed criteria to evaluate the suitability of free tools and public data in incident response practices of national CSIRTs – Towards RQ4

The key findings from Chapter 6 are discussed below. These findings derive from semi-structured interviews and evaluation exercises conducted around two candidate tools and a data source to validate the candidate criteria. The findings are towards answering RQ4.

The validity of the candidate criteria identified in Chapter 5 of this thesis is unproven. Hence, this study concentrated on validating the aforementioned candidate criteria for usefulness, deployment and applicability in practice. The key findings include insights from semi-structured interviews and evaluation exercises with two candidate tools and a data source, as discussed below.

**Perception on the usefulness and deployment of the candidate criteria in national CSIRTs – findings from the semi-structured interviews** As shown in Section 6.3.1, enlightening perceptions were received from the participants concerning the usefulness of the candidate criteria. This includes the completeness and comprehensiveness of the candidate criteria. Such positive and impressive opinions about the candidate criteria show that national CSIRTs are open to new ideas and suggestions. Furthermore, the opinions and views of participants validate the candidate criteria for usefulness, answering the RQ. Interestingly, participants praised the method used in this study to construct and then validate the candidate criteria. Interestingly, participants praised the method used in this study to construct and then validate the candidate criteria, demonstrating their receptive attitude towards this research and its ability to enhance their operational practices. Simultaneously, this encourages more research activities related to national CSIRT operations.

It is important to highlight that some participating national CSIRTs have their own criteria for evaluating tools and data, exemplified by NCSC-FI and CERT-INCIBE. These criteria are just a smaller subset of this study's candidate criteria, often with "Quality in Use" criteria missing, making it less comprehensive. This proves that a complete set of criteria for evaluating free tools and public data is lacking in the participating national CSIRTs. Still, it can be enhanced by consolidating the criteria constructed in this research. This finding is coherent with findings reported in Chapter 4 concerning the lack of systematic procedures for evaluating free tools and public data in national CSIRTs.

Notably, the participants' constructive feedback regarding this research provides further evidence for more proactive research on enhancing the operations of national CSIRTs. This is because there is a general lack of systematic treatment and rigour in how decisions are made within national CSIRTs. For instance, this is reflected in one participant's comment below, which could be eye-opening for national CSIRTs and the research community.

> "*Research should probably be proactive about this kind of issue, and I think, in that sense, I think that your approach will be helpful.*" (Interviewee, JpCERT/CC)

The optimistic feedback from the participants on the deployment of the candidate criteria gave the impression that the candidate criteria are ready for deployment in national CSIRTs. One representative comment from a participant is given below:

> "*There has to be a formal way of doing something. It should, even if it is a SOC, even if it is a CERT that it has a lot of it is deals with a lot of many constituencies, many stakeholders. There has to be a baseline at least.*" (Interviewee, Uganda CERT)

It is particularly encouraging to see most participants are willing to recommend the criteria to other national CSIRTs for deployment. This implies the confidence and trust level in the candidate criteria among the participants, which could also benefit the broader national CSIRT community. This also shows that participants are supportive and optimistic of the criteria derived from this research. This is by considering the value and flexibility of the criteria for the national CSIRT community and how others can tailor them to meet their local needs. One comment on this aspect is given below:

> "*But I mean, yeah, I mean, if there's any, like any sort of like criteria or recommendations for national CSIRT, I can definitely show them, you know, hey, this is something that they're using for the national CSIRT community. So I mean, I don't really ask them to comply completely, but they can refer to it.*" (Interviewee, JpCERT/CC)

Echoing JpCERT's comment, one participant suggested that national CSIRTs could also refer to the criteria when planning to develop software. This implies the applicability of the criteria constructed from this research is twofold: to evaluate free tools and public data and to guide software development. This is as commented below:

> "*this was also a very important for us when we decided to develop a new tool.*"
> (Interviewee, NCSC-Switzerland)

**Applicability of the candidate criteria in practice to evaluate free tools and public data – findings from the tools and data evaluation exercises.** Although interview participants gave positive feedback about the criteria, applying the candidate criteria to concrete tools and data sources to gain more direct evidence on the usefulness, deployment, and critical hands-on tips is equally essential. The evaluation exercises with two candidate tools and a data source validated the criteria on real-world applications to evaluate free tools and public data.

The evaluation of VirusTotal and Hybrid Analysis tools with the criteria informed the differences between the two tools. During the evaluation exercise, the differences between VirusTotal and Hybrid Analysis were observed in terms of "User interface aesthetics", "Accessibility", "Performance efficiency", "Interoperability", "Learnability" and "Supportability". Hence, this indicates the potential ability of the criteria to distinguish between two tools. This could highlight the usability of one tool over the other and potentially help in the decision-making of suitable tools to support incident responses. Moreover, it helps to provide a more systematic way of identifying suitable tools to support incident response by evaluating them with a set of criteria, as opposed to current practices reported in Chapter 4 and Chapter 5. This implies that selecting suitable tools to help national CSIRTs' daily incident response could be done systematically. Nevertheless, it should be noted that the purpose of evaluating two candidate tools in this research is not to highlight that the criteria can be used to find the best tool but to distinguish the differences between them. The differences identified could help national CSIRTs decide on a tool that suits their needs and requirements.

It is also noteworthy that applying the candidate criteria to evaluate two candidate tools and one candidate data source indicates that special care is needed when considering the relevancy of each criterion and what metric(s) could be defined to capture more concrete requirements of each criterion. This is not trivial and requires some guidelines and good case studies as examples to inform staff of national CSIRTs on how the general criteria can be contextualised for different tools and data sources. It is recommended that the evaluation is best done by a team rather than a single staff member to establish a consensus and avoid misjudgements.

Furthermore, the evaluation exercises showed that it was fairly less complicated to translate "Product Quality" criteria (security, usability, maintainability, compatibility, functionality, performance efficiency, reliability, compliance and certification) into concrete metrics since they are primarily about static and factual properties of the tool or data source evaluated. In comparison, it was generally more complicated to consider how to handle "Quality in Use" criteria (context coverage, usability, freedom from risk and popularity) since they are often about end users' opinions and more subjective judgements on the evaluated tool and data source.

Hence, during the evaluation exercise, the researcher played the role of a staff member of a national CSIRT, and the values of the metrics were determined according to the researcher's personal opinions as a staff member of a national CSIRT. However, when deploying the criteria to evaluate tools and data in real-world practices, it is suggested that a team of evaluators should be created, and an operational procedure should be set up to ensure the evaluation results are not biased due to a single staff member's view. This points to some iterations and an auditing step after the evaluation.

## 7.2 Summary

This chapter summarised the key findings of this research as reported in Chapters 3–6. In conclusion, the limitations of the research, including the shortcomings encountered while conducting the research and future work, will be presented in the next chapter.

# Chapter 8

# Conclusion, Limitations and Future Work

This chapter concludes the research by mapping the research conducted to the respective RQs, explained in Section 8.1, achieving the aim of this research. The chapter also summarises the original contributions of this research to knowledge and practice described in Section 8.2. Finally, the chapter provides recommendations for prospective future work as a way forward to continue and bring this research to a higher level described in Section 8.4.

## 8.1 Mapping the Research Conducted to the Respective RQs

Throughout the thesis, pertinent research works and their findings have been presented. This section explains the mapping of the research conducted to the respective RQs as below:

RQ1 *What is the state of the art regarding the use of free tools and public data in the operations of national CSIRTs?*

A systematic literature review (SLR) was undertaken and presented in Chapter 3 – mapped to RQ1. The SLR provided general insights into the operational practices of national CSIRTs regarding the use of free tools and public data from the research literature, national CSIRTs and cross-CSIRTs websites, hence answering RQ1. The SLR demonstrated factual information from the literature, national CSIRTs, and cross-CSIRTs' websites concerning national CSIRTs' operational practices regarding free tools and public data. Despite this, the SLR found that the information provided in the research literature, national CSIRTs, and cross-CSIRT websites is limited, fragmented, and incomplete, which is a research gap. Hence, it informed the need to conduct more empirical research to gain comprehensive and insightful information on the operational practices of national CSIRTs regarding the use of free tools and public data.

RQ2 *How are free tools and public data used in the real-world operations of national CSIRTs?*

An empirical study into the real-world operational practices of national CSIRTs regarding the use of free tools and public data was undertaken and presented in Chapter 4 – mapped to RQ2. The study was conducted at a single national CSIRT – MyCERT and 12 other national CSIRTs worldwide. The empirical research provided insights and a better understanding of how free tools, OSINT, and public data are used in the participating national CSIRTs to support incident response. Furthermore, staff members perceived free tools and public data as useful in the operations of national CSIRTs. Notably, OSINT was introduced as it gained interest in this study. The study confirmed the ad-hoc use of free tools, including OSINT and public data, in national CSIRTs. Moreover, it lacked a systematic approach and procedure to validating such tools and data for quality and usability – an operational gap. This shall inform constructing procedures to enhance current practices concerning free tools and public data. One such enhancement concerns developing criteria for systematically evaluating free tools and public data.

RQ3 *What criteria can be used to evaluate the suitability of free tools and public data in incident response practices of national CSIRTs?*

An empirical study with seven focus groups comprising 15 national CSIRTs was undertaken and presented in Chapter 5 – mapped to RQ3. This is to understand how free tools and public data are evaluated in national CSIRTs, which inform the construction of a set of candidate criteria. First, the study provided insights into the practices of how free tools and public data are evaluated in national CSIRTs. Then, this informed the construction of a set of candidate criteria for evaluating free tools and public data from the opinions of staff members of national CSIRTs to guide a more systematic evaluation for national CSIRTs. Nevertheless, the candidate criteria's usefulness, deployment and applicability are uncertain and not proven. Therefore, an empirical study is suggested to validate the candidate criteria before implementing them in national CSIRT operations.

RQ4 *How can a proposed set of criteria help to evaluate the suitability of free tools and public data in incident response practices of national CSIRTs?*

An empirical study to validate the aforementioned candidate criteria on how the candidate criteria are perceived and the applicability of the criteria is performed and presented in Chapter 6 – mapped to RQ4. First, semi-structured interviews were conducted to gain insights into how participants perceived the usefulness and deployment of the candidate criteria for evaluating free tools and public data. Next, the criteria were validated more objectively with two candidate tools and a data source for applicability in practice. The study found that staff members perceived the candidate criteria as useful and could be deployed in national CSIRTs. The evaluation exercises confirmed the applicability of the criteria in practice to evaluate free tools and public data.

The original contribution of the research demonstrates the findings from the research work mapped to the RQs, explained in the next section below:

## 8.2 Summary of Research Original Contributions

The research made significant contributions to knowledge and practice. The findings relating to the ad-hoc practices in the operations of national CSIRTs, the use of free tools and public data to respond to cyber incidents and, more importantly, the validated criteria for evaluating free tools and public data – are contributions to knowledge. The validated criteria are also a pivotal contribution to practice as national CSIRTs and CSIRTs, in general, could now evaluate their toolsets systematically for quality and usability purposes.

### 8.2.1 Empirical evidence confirming that national CSIRTs use free tools and public data in responding to cyber security incidents. This contribution is towards answering RQ2.

The research confirmed and provided real-world empirical evidence that national CSIRTs use free tools and public data to support incident responses in their operational practices. This contribution is towards answering RQ2. The evidence is original and contributes to the existing body of knowledge, providing a better understanding of the operational practices in real-world national CSIRTs and addressing a gap in the research. The free tools, including open-source tools, free OSINT tools and public data used in real-world national CSIRTs to support incident responses, were identified in the research. The research participants described the purpose and functions of each free tool and public data, and they are compiled and incorporated into this thesis. This compilation can be a significant reference for practitioners looking for reliable free tools and public data to support incident response, besides expanding the knowledge about free tools and public data used in national CSIRTs. Empirical evidence of how the participating national CSIRTs perceived free tools, OSINT tools and public data was revealed in this research. Participants perceived free tools, including open-source and OSINT tools, as useful, valuable and beneficial for incident response in national CSIRTs. Such perceptions are crucial to confirm that free tools and public data are useful in supporting incident responses.

### 8.2.2 Confirmation through empirical evidence that the use of free tools and public data by national CSIRTs is ad-hoc and not governed by formal, defined and institutionalised procedures. This contribution is towards answering RQ2.

The research confirmed with empirical evidence that free tools and public data are used ad-hoc without systematic and institutionalised procedures in the surveyed and interviewed national CSIRTs reported in Chapter 4 – one use is on the evaluation of free tools and public data. This contribution is towards answering RQ2. Furthermore, a focus group study reported in Chapter 5 substantiated this evidence, revealing in detail the informal approaches currently used in national CSIRTs to evaluate free tools and public data. The finding is original and contributes to the existing body of knowledge, providing insights into how free tools and public data are evaluated in national CSIRTs. The finding is significant because it informs researchers and stakeholders to develop more systematic procedures for national CSIRTs.

### 8.2.3 A set of criteria to evaluate relevant attributes of free tools and public data. This contribution is towards answering RQ3.

The research identifies 14 criteria for evaluating free tools and eight criteria for evaluating public data from the collective opinions of research participants reported in Chapter 5. This contribution is towards answering RQ3. The criteria are original and contribute to the existing body of knowledge, providing insights into a set of candidate criteria for evaluating free tools and public data. The criteria are significant as they were constructed from collective opinions and views of staff members of national CSIRTs worldwide in a focus group study.

The candidate criteria for evaluating tools are divided into two categories – “Product Quality” and “Quality in Use” as per the ISO/IEC 25010:2011 standard. The candidate criteria concentrated on “Product Quality” with their definitions (descriptions) is shown in Table 20. The candidate criteria focused on “Quality in Use” with their definitions (descriptions) are shown in Table 21. The candidate criteria for evaluating data with their definitions (descriptions) are shown in Table 22.

### 8.2.4 Validation of the applicability of the aforementioned criteria in evaluating free tools and public data in national CSIRTs. This contribution is towards answering RQ4.

The research contributes to the validation of the applicability of the aforementioned criteria. This contribution is towards RQ4. The validation was done in two ways: 1) the opinions of staff members of national CSIRTs about how they perceive the usefulness and deployment of the criteria in national CSIRTs, and 2) more objective validation of the criteria for applicability using several metrics to evaluate two candidate tools and a data source. Participants' perceptions confirmed the usefulness and deployment of the criteria to evaluate free tools and public data in national CSIRTs. These perceptions can be referred to in Chapter 6. Applying the criteria to evaluate two candidate tools and a data source confirms the applicability and usability of the criteria in practice to evaluate free tools and public data. The evaluation result with two sample tools and a data source can be viewed at Appendixes A and B. Significantly, the validated criteria can be used to measure the maturity of toolsets for national CSIRTs and CSIRTs with potential applications of the criteria in broader security operations.

Despite the contributions, more prospective future work could be undertaken to expand the research reported in this thesis to a higher level, as suggested in the Future Work Section.

## 8.3 Limitations

This section discusses the limitations of the research. Some of these limitations were reported in previous studies related to information security [244; 138].

**Data collection – number, diversity and selection of participants.** First, the number of participants and the number of national CSIRTs covered in this research, reported in Chapters 4, 5 and 6 may be relatively low, which might affect the generalisability of the results reported in the research [101]. Nevertheless, it should be noted that for a very closed community like national CSIRTs, recruiting a small number of willing and qualified participants from a small community like national CSIRTs was not easy. Although the research attempted to recruit more participants, it proved challenging. This was not unexpected given the target pool of participants is very niche and the small community of national CSIRTs. Such difficulties have also been reported in past studies [31; 189; 200].

Despite the small number of participants, the key findings of the study are valid and reliable considering the following two facts: 1) the participants' opinions are highly consistent among one another, and there is a consensus on the RQ being investigated, and the saturation effect is already observed even with such a small sample [72]; indicating further data collection is unnecessary [223]; 2) the key findings also match the researcher's experience as an employee of a national CSIRT (MyCERT) for over 20 years, 3) the key findings are coherent with informal feedback received via different channels such as conferences and seminars, from staff members of MyCERT and other national CSIRTs and 4) notably this is a qualitative research and not quantitative or statistical study; hence number is not a significant concern in this research. Nevertheless, this research acknowledges the limitation and suggests conducting follow-up research with a more significant number of participants representing a broader range of national CSIRTs.

A related limitation is the diversity of participants – the sampling for the study may lack diversity in terms of participants' experiences working in national CSIRTs, which can constrain the depth and breadth of insights gained. The researcher mitigated this by using probing questions, clear, simple language and open-ended interviews to encourage participants, especially less experienced participants, to think critically and provide more detailed responses. This limitation was also mitigated, considering feedback from less-experienced participants is consistent with the more experienced participants. Furthermore, the feedback is consistent and matches with the researcher's experience of more than 20 years as a staff member of a national CSIRT.

Another limitation related to data collection is researcher bias in the selection of participants. This limitation should account for the difficulty recruiting national CSIRTs to participate in this research. Furthermore, the research was conducted during the global COVID-19 pandemic, which impacted many activities worldwide, including research between 2020 and 2022. This has also affected data collection for this research to some extent. Hence, the researcher used her professional networking with some national CSIRTs, specifically the Malaysia national CSIRT (MyCERT), to mitigate difficulty in recruitment. Despite this, it is also noteworthy to mention that the researcher is constantly reflective of researcher bias, which resulted in the data collection choices and mitigated any impacts on the research due to the choices in data collection. This includes using peer debriefing to summarise the focus group discussions, peer reviews with the research community and journal peer reviews on the research findings. Notably, the journal peer reviews also led to the publication of some research works and findings reported in this thesis in scientific journals.

**Tools and data evaluation exercises.** A second limitation is that the evaluation exercise with two candidate tools and a data source is still relatively limited. Due to time constraints, the study only evaluated two candidate tools and one data source. Furthermore, the researcher did the evaluation manually without assistance from automated tools or software, making the evaluation exercise time-consuming. Automating some of the evaluation tasks would be helpful and save time, potentially enabling the evaluation of more tools. Evaluating more tools could allow a broader range of criteria, as not all criteria are necessarily relevant to specific tools, such as “confidentiality” and “analysability”. This is an observation from the researcher’s evaluation exercise. Nonetheless, it should be noted that the purpose of the evaluation was to gain more direct evidence on how the criteria can be contextualised to evaluate tools and data; hence, the exercises were conducted less formally and involved limited candidate tools and data. Despite this, the study has provided some initial work to contextualise the criteria to evaluate two candidate tools and one data source.

Furthermore, the evaluation exercise in this research is limited insofar as the researcher conducted it on behalf of staff members of national CSIRTs. This could make the evaluation results subject to the researcher’s preference and experience. Furthermore, the researcher is an experienced staff member of a national CSIRT, hence excluding novice users’ perspectives. Moreover, the evaluation was not run under an actual operational environment; therefore, a wider range of users’ perspectives may not be included. Thus, future work could help address this limitation by conducting evaluation exercises with actual users, ranging from novices to experts and staff members of national CSIRTs.

A related limitation is evaluating the “Quality in Use” criteria for tools, which primarily focuses on end-user satisfaction. This would require a more comprehensive evaluation, potentially a qualitative or quantitative (statistical method) approach. This could be using interviews or surveys with staff members of national CSIRTs as the respondents to obtain their opinions and satisfaction levels regarding using the tools. For example, the criteria “User Experience”, “Flexibility” and “Usefulness” would be best to obtain the opinions from staff members of national CSIRTs to yield more complete results. In this research, evaluations using interviews or surveys could not be performed due to time constraints, restricting the depth and breadth of the evaluation exercise. The COVID-19 pandemic further compounded the limitations regarding time and other resources, i.e. budget. Therefore, this might limit the completeness of the evaluation exercise results presented in this research.

A minor limitation of the criteria is their high-level nature; therefore, translating them into more concrete metrics and values is not trivial. Hence, applying them to evaluate specific tools and data is challenging. Some difficulties include more time needed to understand and translate the criteria into relevant metrics. This was one of the key observations during the evaluation exercise reported in Section 6.3.2. Ample time is also required to study and understand the tools and data that will be evaluated. Skills and expertise in conducting evaluation exercises are another limitation encountered. Such skills and expertise are essential to guide national CSIRTs in identifying concrete metrics and making good judgement during the evaluation.

Considerable investment in translating the criteria is required before they can be fully used. During the evaluation, it was also challenging to consider the relevancy of each criterion and what metric(s) could be defined to capture more concrete requirement(s) of each criterion. This would require more work on the criterion-to-metric process and the creation of more detailed guidelines and case studies. The guidelines and case studies will be very helpful in guiding national CSIRTs in conducting the evaluation exercise in real-world operational practices.

**Coding.** A third limitation encountered while conducting the research is maintaining the intra-coder reliability. This refers to consistency in how the same person codes data at multiple time points [196]. If the same person returns to the same data at another time, how similar would the output be today to the work from before? It's unlikely to be an exact match. It's fairly common for a drift in coding during a task like this (at the first item, the coder has a different concept in mind for codes than at the last item). Several factors could affect this, such as fatigue due to coding large volumes of data and returning to code the same data after a long time interval due to an illness.

To overcome the intra-coder limitation in this research, the researcher reviewed the codes multiple times and iterated them before coming up with a final list. However, it still remains a time-consuming task and when planning further research, enhanced training and guidance for coders on coding tasks may help address the limitation. Furthermore, member-checking or participant validation of the codes and journal peer-reviews may help to address the intra-coder limitation [40]. Additionally, using computer-aided tools to manage the coding could help address the intra-coder limitation. Such tools can keep track of the coding decisions, notes, and memos at each stage. This information is retrievable later, helping to improve the coding consistency.

**Time.** A fourth limitation encountered while conducting this research is time. As the PhD research is to be accomplished within a stipulated time by the institution and complying with the funder's requirements, it became a challenge to accomplish this research within the specified time frame. Much of the time for this research was spent on data collection to ensure sufficient data was collected to align with the research's scope and answer the RQs. The unprecedented COVID-19 Pandemic further compounded the limitation in time, primarily causing delays and slowness in the recruitment of participants.

Much time was also spent on data analysis of qualitative data, i.e., interviews and focus group data, which was time-consuming and cumbersome. This included manually transcribing all the interviews and focus group recordings to ensure the accuracy of the transcriptions, which was challenging. A huge amount of time was also spent on coding. As discussed, this was to ensure the reliability of the coding, specifically intra-coder reliability and the overall rigour of the research. Generally, time constraints and complexity in analysing qualitative data are acknowledged in qualitative research. This limitation can be addressed by limiting the scope of the research by focusing on a more precise and specific topic. Broader research topics may be prone to generating unmanageable volumes of data, making data analysis a complex and time-consuming task.

Having more time could have enabled more empirical studies related to this research. One such study could be a comparative study between matured national CSIRTs and less-matured national CSIRTs. Such a study would help identify the gaps between matured and less-matured national CSIRTs and identify measures to bridge the gaps.

Overall, in addressing the limitations of this research, first of all, the researcher set a boundary for the research in terms of the scope and aim of the research. Next, tasks were prioritised, processes streamlined, and resource usage was planned to avoid crossing the boundary set for this research. More importantly, the researcher acknowledged and communicated these limitations to ensure the transparency of the research and its findings. Nevertheless, this study provides opportunities for further research in the immediate future and is suggested in the next section. Future works could help address some of the limitations identified in this research.

## 8.4 Future Work

This research makes several recommendations on how other researchers can build upon and expand the present research by taking the research further to ensure the continuity of research in this field. This includes further developing the body of knowledge in this research and enhancing practices with more systematic procedures. Interestingly, some future work recommended here could address the limitations of this research reported in Chapter 7.

The recommendations for future work are outlined below:

1. Include more participants from a broader range of national CSIRTs and diverse demographics and ensure a non-biased selection of participants. Prospective future work in this field of study with more participants representing wider national CSIRTs is recommended to obtain a broader range of opinions and experiences from many national CSIRTs, potentially addressing the limitation in the number of participants. Notably, future work must consider the selection of participants, and national CSIRTs must be independent of researcher bias. Participants from more diverse demographics are suggested for future work. Future work could help validate the findings presented in this research and take note of any significant differences in the new findings. It is also recommended for future work to include participants from non-national CSIRTs to obtain views and experiences that may differ. These could be academic, organisational, and enterprise/industrial CSIRTs. This may enable CSIRTs of various types to share their best practices and learn from one another, overall improving incident response operations across numerous

sectors and domains.

2. Expand the research scope to evaluate more candidate tools and data sources. Future work is suggested to expand the evaluation exercises with more candidate tools and data sources. The exercises in this research were performed less formally since the purpose was to gain preliminary evidence on how the candidate criteria can be used and contextualised. More formal and comprehensive evaluation exercises involving a greater range of candidate tools and data sources could help consolidate the results reported in this thesis. These should be evaluated against a greater set of staff through using surveys or interviews. This will complement the exercises conducted in this research, especially on the "Quality in Use" criteria, which rests on more subjective values.

   Additionally, future work is suggested on the criterion-to-metric process and creating more detailed guidelines and case studies. Such future work would be helpful as the candidate criteria are largely high-level; translating them into more concrete metrics and values can be challenging. Nonetheless, these guidelines and case studies would be helpful to various CSIRTs when evaluating tools and data in their operations. Future research could supported with Artificial Intelligence (AI) based techniques, tool automation, and data evaluation. The research reported in this thesis was done manually, yet there are multiple areas where AI-based mechanisms will help evaluate tools and data more efficiently, objectively, accurately and reproducibly.

3. Developing a taxonomy of tools and data sources with metrics that connect to the criteria. Future works to expand the criteria and metrics from this research to construct a taxonomy or even an ontology that will connect the criteria, different types of tools and data sources used by (national and non-national) CSIRTs is suggested. Doing so will inform the development of more practical operational guidelines and potentially enable partial automation of the tool and data evaluation procedure. In addition, more metrics and scoring systems can be identified and tested for even more quantitative and reproducible evaluation exercises.

4. In-depth study of the following criteria: Compliance, Popularity and Certification. Future work is suggested for conducting a more in-depth analysis of the three criteria for tool evaluation identified from this research and unavailable in the ISO/IEC 25000 SQuaRE Model, which is *Compliance, Popularity and Certification.* These three criteria could be merged into existing criteria and reflected as separate metric(s) under another criterion. The criterion could also be relevant for data evaluation, although this was not considered in this research. If confirmed that these criteria should remain separate without merging, it will be helpful to work with the

international standardisation community to add them to future editions of relevant international standards, especially ISO/IEC 25010 and ISO/IEC 25012.

5. In-depth study on various theories that were not considered to understand or explain the phenomena observed in this research. For example, the Technology Adoption Model (TAM) [270; 127] may help explain why national CSIRTs adopt and use free tools and public data and their perceived usefulness of free tools and public data. Furthermore, the Institutional Theory may be relevant, given the observation from this research concerning the lack of institutionalised procedures for adopting free tools and public data in the participating national CSIRTs. Other theories that might be relevant and interesting for future work include the Diffusion of Innovations Theory [53; 152] – to study the adoption and diffusion of free tools, and the Organisational Learning Theory [3; 174] – to analyse how CSIRTs adapt and evolve their practices over time.

# Bibliography

[1] Philip Adu. 2019. *A Step-by-Step Guide to Qualitative Data Coding.* Routledge.

[2] Arezoo Aghaei Chadegani, Hadi Salehi, Melor Yunus, Hadi Farhadi, Masood Fooladi, Maryam Farhadi, and Nader Ale Ebrahim. 2013. A Comparison between Two Main Academic Literature Collections: Web of Science and Scopus Databases. *Asian social science* 9, 5 (2013), 18–26. https://doi.org/10.5539/ass.v9n5p18

[3] Atif Ahmad, Kevin C Desouza, Sean B Maynard, Humza Naseer, and Richard L Baskerville. 2020. How integration of cyber security management and incident response enables organizational learning. *Journal of the Association for Information Science and Technology* 71, 8 (2020), 939–953.

[4] Atif Ahmad, Justin Hadgkiss, and A. B. Ruighaver. 2012. Incident Response Teams – Challenges in Supporting the Organisational Security Function. *Computers & Security* 31, 5 (2012), 643–652. https://doi.org/10.1016/j.cose.2012.04.001

[5] Atif Ahmad, Sean B. Maynard, Kevin C. Desouza, James Kotsias, Monica T. Whitty, and Richard L. Baskerville. 2021. How can Organizations Develop Situation Awareness for Incident Response: A Case Study of Management Practice. *Computers & Security* 101 (2021), 102122. https://doi.org/10.1016/j.cose.2020.102122

[6] Atif Ahmad, Sean B. Maynard, and Graeme Shanks. 2015. A Case Analysis of Information Systems and Security Incident Responses. *International Journal of Information Management* 35, 6 (2015), 717–723. https://doi.org/10.1016/j.ijinfomgt.2015.08.001

[7] Nisar Ahmad, Angeliki N. Menegaki, and Saeed Al-Muharrami. 2020. Systematic literature review of tourism growth nexus: An overview of the literature and a content analysis of 100 most influential papers. *Journal of Economic Surveys* 34, 5 (2020), 1068–1110. https://doi.org/10.1111/joes.12386

[8] Rabiah Ahmad, Zahri Yunos, Shahrin Sahib, and Mariana Yusoff. 2012. Perception on Cyber Terrorism: A Focus Group Discussion Approach. *Journal of Information Security* 3 (2012), 231–237. https://doi.org/10.4236/jis.2012.33029

[9] Rahayu Azlina Ahmad and Mohd Shamir Hashim. 2011. The Organisation of Islamic Conference–Computer Emergency Response Team(OIC-CERT): Answering Cross Border Cooperation. In *Proceedings of the 2011 2nd Worldwide Cybersecurity Summit*. IEEE, 5 pages. https://ieeexplore.ieee.org/document/5978783

[10] Anas Bassam Al Badareen, Mohd Hasan Selamat, Marzanah A. Jabar, Jamilah Din, and Sherzod Turaev. 2011. Software Quality models: A Comparative Study. In *Software Engineering and Computer Systems, Part I: Second International Conference, ICSECS 2011, Kuantan, Malaysia, June 27-29, 2011. Proceedings, Part I*. Springer, 46–55. https://doi.org/10.1007/978-3-642-22170-5_4

[11] Rafa E. Al-Qutaish. 2010. Quality Models in Software Engineering Literature: An Analytical and Comparative Study. *Journal of American Science* 6, 3 (2010), 166–175. http://www.jofamericanscience.org/journals/am-sci/am0603/22_2208_Qutaish_am0603_166_175.pdf

[12] Jose Antonio Mulet Alberola and Irene Fassi. 2022. Towards the Assessment of Performance-based Interactions in Collaborative CPPS. *Procedia Computer Science* 200 (2022), 1636–1645. https://doi.org/10.1016/j.procs.2022.01.364

[13] Steve Alder. 2021. Cyber Incident Notification of Act 2021. web page. https://www.hipaajournal.com/cyber-incident-notification-act-of-2021-introduced-in-the-senate/

[14] APCERT. 2021. *APCERT Annual Report 2021*. Annual report. APCERT. http://www.apcert.org/documents/pdf/APCERT_Annual_Report_2021.pdf

[15] William A. Arbaugh, William L. Fithen, and John McHugh. 2000. Windows of Vulnerability: A Case Study Analysis. *Computer* 33, 12 (2000), 52–59. https://doi.org/10.1109/2.889093

[16] Australian Computer Emergency Response Team AusCERT. 2023. CAustralia's leading cyber emergency response team. webpage. https://www.auscert.org.au/

[17] Gayane Azizyan, Miganoush Katrin Magarian, and Mira Kajko-Matsson. 2011. Survey of Agile Tool Usage and Needs. In *Proceedings of the 2011 Agile Conference*. IEEE, 29–38. https://doi.org/10.1109/AGILE.2011.30

[18] Riza Azmi, William Tibben, and Khin Than Win. 2016. Motives behind Cyber Security Strategy Development: A Literature Review of National Cyber Security Strategy. , 12 pages. https://aisel.aisnet.org/acis2016/52/

[19] Shahriar Badsha, Iman Vakilinia, and Shamik Sengupta. 2019. Privacy Preserving Cyber Threat Information Sharing and Learning for Cyber Defense. In *Proceedings of the 2019 IEEE 9th Annual Computing and Communication Workshop and Conference*. IEEE, 708–714. https://doi.org/10.1109/CCWC.2019.8666477

[20] Ibrahim Baggili and Marcus Rogers. 2009. Self-reported Cyber Crime: An Analysis on the Effects of Anonymity and Pre-employment Integrity. *International Journal of Cyber Criminology* (2009). https://digitalcommons.newhaven.edu/electricalcomputerengineering-facpubs/26/

[21] Helen L. Ball. 2019. Conducting online surveys. *Journal of Human Lactation* 35, 3 (2019), 413–417. https://doi.org/10.1177/0890334419848734

[22] Rosaline Barbour. 2008. *Doing Focus Groups*. SAGE. https://doi.org/10.4135/9781849208956

[23] Nicole Basaraba. 2021. A bottom-up Method for Remixing Narratives for Virtual Heritage Experiences. *Convergence: The International Journal of Research into New Media Technologies* 28 (2021). Issue 6. https://doi.org/10.1177/13548565211048

[24] Tom Bateman. 2022. Russia Led Major Cyberattack on European Broadband Network just before Ukraine Invasion, says West. Web page. https://www.euronews.com/next/2022/05/10/west-says-russia-led-major-cyber-attack-on-satellite-broadband-network-just-before-ukraine

[25] Mariette Bengtsson. 2016. How to Plan and Perform a Qualitative Study Using Content Analysis. *NursingPlus Open* 2 (2016), 8–14. https://doi.org/10.1016/j.npls.2016.01.001

[26] Nigel Bevan. 2001. International Standards for HCI and Usability. *International Journal of Human-Computer Studies* 55, 4 (2001), 533–552. https://doi.org/10.1006/ijhc.2001.0483

[27] Tracy Bills, Brittany Manley, and James Lord. 2022. *Enabling the Sustainability and Success of a National Computer Security Incident Response Team*. Technical Report. Carnegie Mellon University. https://resources.sei.cmu.edu/asset_files/Handbook/2022_002_001_885865.pdf

[28] Barry W. Boehm, John R. Brown, and Mlity Lipow. 1976. Quantitative Evaluation of Software Quality. In *Proceedings of the 2nd International Conference on Software Engineering*. ACM, 592–605. https://doi.org/10.5555/800253.807736

[29] Germinal Boloix and Pierre N. Robillard. 1995. A Software System Evaluation Framework. *Computer* 28, 12 (1995), 17–26. https://doi.org/10.1109/2.476196

[30] Belinda Ivy Botchway, Akinwonmi Akintoba Emmanuel, Nunoo Solomon, and Alese Boniface Kayode. 2021. Evaluating Software Quality Attributes using Analytic Hierarchy Process (AHP). *International Journal of Advanced Computer Science and Applications* 12, 3 (2021), 165–173. https://doi.org/10.14569/IJACSA.2021.0120321

[31] David Botta, Rodrigo Werlinger, André Gagné, Konstantin Beznosov, Lee Iverson, Sidney Fels, and Brian Fisher. 2007. Towards Understanding IT Security Professionals and Their Tools. In *Proceedings of the 3rd Symposium on Usable Privacy and Security*. ACM, 100–111. https://doi.org/10.1145/1280680.1280693

[32] Romain Bourgue, Joshua Budd, Jachym Homola, Michal Wlasenko, and Dariusz Kulawik. 2013. *Detect, SHARE, Protect: Solutions for Improving Threat Data Exchange among CERTs*. Technical Report. https://www.enisa.europa.eu/publications/detect-share-protect-solutions-for-improving-threat-data-exchange-among-certs

[33] Virginia Braun and Victoria Clarke. 2006. Using Thematic Analysis in Psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[34] Chuck Brooks. 2022. Alarming Cyber Statistics For Mid-Year 2022 That You Need To Know. Web page. https://www.forbes.com/sites/chuckbrooks/2022/06/03/alarming-cyber-statistics-for-mid-year-2022-that-you-need-to-know/?sh=2c10d2b87864

[35] Nevil Brownlee and Erik Guttman. 1998. Expectations for Computer Security Incident Response. Web page. https://www.ietf.org/rfc/rfc2350.txt

[36] Nevil Brownlee and Erik Guttman. 2016. Directive (EU) 2016/1148 of the European Parliament and of the Council. Web page. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2016.194.01.0001.01.ENG&toc=OJ:L:2016:194:TOC

[37] Mario Callegaro, Katja Lozar Manfreda, and Vasja Vehovar. 2015. *Web Survey Methodology*. SAGE. https://doi.org/doi.org/10.4135/9781529799651

[38] Philip Cash and Chris Snider. 2014. Investigating Design: A Comparison of Manifest and Latent Approaches. *Design Studies* 35, 5 (2014), 441–472. https://doi.org/10.1016/j.destud.2014.02.005

[39] Catherine Cassell. 2015. *Conducting Research Interviews for Business and Management Students*. Sage. https://doi.org/10.4135/9781529716726

[40] Ashley Castleberry and Amanda Nolen. 2018. Thematic analysis of qualitative research data: Is it as easy as it sounds? *Currents in pharmacy teaching and learning* 10, 6 (2018), 807–815.

[41] Stephen Cavanagh. 1997. Content analysis: Concepts, Methods and Applications. *Nurse Researcher* 4, 3 (1997), 5–16. https://doi.org/10.7748/nr.4.3.5.s2

[42] Paul Cichonski, Tom Millar, Tim Grance, and Karen Scarfone. 2012. *Computer Security Incident Handling Guide*. Technical Report 800-61 Revision 2. National Institute of Standards and Technology, U.S. Department of Commerce. https://doi.org/10.6028/NIST.SP.800-61r2

[43] Felipe Sabino Costa. 2021. *A Practical Approach to Adopting the IEC 62443 Standards*. White paper. MOXA, Inc. https://www.moxa.com/en/literature-library/moxa-a-practical-approach-to-adopting-the-iec-62443-standards-white-paper

[44] PCI Security Standards Council. 2018. Payment Card Industry (PCI) Data Security Standard. web page. (2018). https://www.pcisecuritystandards.org/about_us/

[45] John W. Creswell and Dana L. Miller. 2000. Determining Validity in Qualitative Inquiry. *Theory Into Practice* 39, 3 (2000), 124–130. https://doi.org/10.1207/s15430421tip3903_2

[46] Carolyn Folkman Curasi. 2001. A Critical Exploration of Face-to-Face Interviewing vs. Computer-mediated Interviewing. *International Journal of Market Research* 43, 4 (2001), 1–13. https://doi.org/10.1177/147078530104300402

[47] Cyber Security Agency of Singapore. 2021. Singapore Cyber Security Strategy. Governmental report. https://www.csa.gov.sg/News/Publications/singapore-cybersecurity-strategy-2021

[48] CyberEdge. 2022. 2021 Cyberthreat Defense Report. Web page. https://www.imperva.com/resources/resource-library/reports/2021-cyberthreat-defense-report/

[49] Cybersecurity and Infrastructure Security Agency CISA. [n.d.]. Cyber Incident Reporting for Critical Infrastructure Act of 2022 CIRCIA. web page. https://www.cisa.gov/circia

[50] Tharashasank Davuluru, Jayapal Medida, and V. S. K. Reddy. 2014. A Study of Software Quality Models. In *Proceedings of the 2014 International Conference on Advances in Engineering & Technology Research*. IEEE, 8 pages. https://doi.org/10.1109/ICAETR.2014.7012958

[51] Tony Day, Helen Gibson, and Steve Ramwell. 2016. Fusion of OSINT and Non-OSINT Data. In *Open Source Intelligence Investigation: From Strategy to Implementation*. Springer, 133–152. https://doi.org/10.1007/978-3-319-47671-1_9

[52] Stephanie de Smale, Rik van Dijk, Xander Bouwman, Jeroen van der Ham, and Michel van Eeten. 2023. No One Drinks From the Firehose: How Organizations Filter and Prioritize Vulnerability Information. In *Proceedings of the 2023 IEEE Symposium on Security and Privacy*. IEEE, 203–219. https://doi.org/10.1109/SP46215.2023.00012

[53] James W Dearing and Jeffrey G Cox. 2018. Diffusion of innovations theory, principles, and practice. *Health affairs* 37, 2 (2018), 183–190.

[54] Jessica T. DeCuir-Gunby, Patricia L. Marshall, and Allison W. McCulloch. 2011. Developing and Using a Codebook for the Analysis of Interview Data: An Example from a Professional Development Research Project. *Field Methods* 23, 2 (2011), 136–155. https://doi.org/10.1177/1525822X103884

[55] Tom Deliens, Peter Clarys, Ilse De Bourdeaudhuij, and Benedicte Deforche. 2014. Determinants of eating behaviour in university students: a qualitative study using focus group discussions. *BMC Public Health* 14, 1 (2014), 12 pages. https://doi.org/10.1186/1471-2458-14-53

[56] DomainTools. [n.d.]. DomainTools. website. https://www.domaintools.com/

[57] R. Geoff Dromey. 1995. A Model for Software Product Quality. *IEEE Transactions on Software Engineering* 21, 2 (1995), 146–162. https://doi.org/10.1109/32.345830

[58] Elisabeth Dubois and Unal Tatar. 2022. Mitigating Global Cyber Risk Through Bridging the National Incident Response Capacity Gap. In *Proceedings of the 17th International Conference on Information Warfare and Security*, Vol. 17. Academic Conferences International Limited, 527–531. https://doi.org/10.34190/iccws.17.1.66

[59] Hanneke Duijnhoven, Tom Van Schie, and Don Stikvoort. 2019. *Stimulating the Development and Maturity Enhancement of National CSIRTs V1.0.* Technical Report. 1–28 pages. https://thegfce.org/wp-content/uploads/2020/05/MaturityFrameworkfornationalCSIRTsv1.0_GFCE.pdf

[60] Hanneke Duijnhoven, Tom van Schie, and Don Stikvoort. 2021. *Stimulating the Development and Maturity Enhancement of National CSIRTs*. TNO Publication. https://repository.tno.nl//islandora/object/uuid:e1ba969e-7ab4-4bd5-83fb-7c029db15265

[61] Christen Erlingsson and Petra Brysiewicz. 2017. A hands-on guide to doing content analysis. *African Journal of Emergency Medicine* 7, 3 (2017), 93–99. https://doi.org/10.1016/j.afjem.2017.08.001

[62] European Commission EU. 2017. CyberExchange. web page. https://ec.europa.eu/inea/en/connecting-europe-facility/cef-telecom/2017-eu-ia-0118

[63] European Commission. 2009. Protecting Europe from Large scale Cyber-attacks and Disruptions: Enhancing Preparedness, Security and Resilience. web page. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=celex:52009DC0149

[64] European Parliament and European Council. 2016. Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union. EU legislation. https://eur-lex.europa.eu/eli/dir/2016/1148/oj

[65] Matthew Field. 2018. WannaCry Cyber-attack Cost the NHS £92m after 19,000 Appointments were Cancelled. Web page. https://www.telegraph.co.uk/technology/2018/10/11/wannacry-cyber-attack-cost-nhs-92m-19000-appointments-cancelled/

[66] Murray J. Fisher and Andrea P. Marshall. 2009. Understanding descriptive statistics. *Australian Critical Care* 22, 2 (2009), 93–97. https://doi.org/10.1016/j.aucc.2008.11.003

[67] Craig S. Fleisher. 2008. Using Open Source Data in Developing Competitive and Marketing Intelligence. *European Journal of Marketing* (2008). https://doi.org/10.1108/03090560810877196

[68] Rachel Flynn, Lauren Albrecht, and Shannon D. Scott. 2018. Two Approaches to Focus Group Data Collection for Qualitative Health Research: Maximizing Resources and Data Quality. *International Journal of Qualitative Methods* 17, 1 (2018), 1–9. https://doi.org/10.1177/1609406917750781

[69] European Union Agency for Cybersecurity (ENISA). 2019. MeliCERTes. web page. https://github.com/melicertes/csp

[70] Fiona E. Fox, Marianne Morris, and Nichola Rumsey. 2007. Doing Synchronous Online Focus Groups With Young People: Methodological Reflections. *Qualitative*

*Health Research* 17, 4 (2007), 539–547. https://doi.org/10.1177/1049732306298754

[71] Steven Furnell, Pete Fischer, and Amanda Finch. 2017. Can't get the staff? The Growing Need for Cyber-security Skills. *Computer Fraud & Security* 2017, 2 (2017), 5–10. https://doi.org/10.1016/S1361-3723(17)30013-1

[72] Patricia I Fusch Ph D and Lawrence R Ness. 2015. Are we there yet? Data saturation in qualitative research. (2015).

[73] Dhananjay Gade. 2013. The evaluation of software quality. https://digitalcommons.unl.edu/imsediss/38/

[74] Pamela Brady Germain and Greta G. Cummings. 2010. The Influence of Nursing Leadership on Nurse Performance: A Systematic Literature Review. *Journal of nursing management* 18, 4 (2010), 425–439. https://doi.org/10.1111/j.1365-2834.2010.01100.x

[75] Anita Gibbs. 1997. Focus Groups. *Social Research Update* 19, 8 (1997), 8 pages. https://sru.soc.surrey.ac.uk/SRU19.html

[76] Lisa M. Given. 2008. *The Sage encyclopedia of qualitative research methods*. SAGE.

[77] Michael Glassman and Min Ju Kang. 2012. Intelligence in the Internet Age: The Emergence and Evolution of Open Source Intelligence (OSINT). *Computers in Human Behavior* 28, 2 (2012), 673–682. https://doi.org/10.1016/j.chb.2011.11.014

[78] Michael J. Glennon. 2012. State-level Cybersecurity. *Policy Review* 171 (2012). https://www.hoover.org/research/state-level-cybersecurity

[79] R. D. Goddard and Peter Villanova. 2006. Designing Surveys and questionnaires for research. In *The Psychology Research Handbook: A Guide for Graduate Students and Research Assistants*. SAGE, 114–125. https://doi.org/10.4135/9781412976626.n8

[80] Alessandra Gorini, Massimo Miglioretti, and Gabriella Pravettoni. 2012. A New Perspective on Blame Culture: an Experimental Study. *Journal of Evaluation in Clinical Practice* 18, 3 (2012), 671–675. https://doi.org/10.1111/j.1365-2753.2012.01831.x

[81] Robert B. Grady. 1992. *Practical Software Metrics for Project Management and Process Improvement*. Prentice-Hall, Inc.

[82] George Grispos, William Bradley Glisson, and Tim Storer. 2015. Security Incident Response Criteria: A Practitioner's Perspective. arXiv preprint arXiv:1508.02526. https://arxiv.org/abs/1508.02526

[83] Bernd Grobauer and Thomas Schreck. 2010. Towards Incident Handling in the Cloud: Challenges and Approaches. In *Proceedings of the 2010 ACM Workshop on Cloud Computing Security Workshop*. 77–86. https://doi.org/10.1145/1866835.1866850

[84] J. C. Grundy and J. G. Hosking. 2001. Software tools. In *Wiley Encyclopedia of Software Engineering*. John Wiley & Sons, Inc. https://doi.org/10.1002/0471028959.sof332

[85] Fernando Gualo, Moisés Rodríguez, Javier Verdugo, Ismael Caballero, and Mario Piattini. 2021. Data Quality Certification Using ISO/IEC 25012: Industrial Experiences. *Journal of Systems and Software* (2021), 17 pages. https://doi.org/10.1016/j.jss.2021.110938

[86] Jaber F. Gubrium and James A. Holstein. 2001. *Handbook of Interview Research: Context and Method*. SAGE. https://doi.org/10.4135/9781412973588

[87] Glenn Gumba, Deborah G. Brosas, and Jessie R. Paragas. 2021. Assessment of SIAS Application Using Software Quality Model. In *Proceedings of the 2021 3rd International Conference on Research and Academic Community Services*. IEEE, 197–202. https://doi.org/10.1109/ICRACOS53680.2021.9701982

[88] Muhammad Haidar, Yudho Giri Sucahyo, Teddy Sukardi, and Arfive Gandhi. 2021. Analysis of CSIRT Services in Facing Cyber Security Challenges in Indonesia. In *2021 4th International Conference on Information and Communications Technology (ICOIACT)*. IEEE, 154–159. https://doi.org/10.1109/ICOIACT53268.2021.9563925

[89] Elizabeth J. Halcomb, Leila Gholizadeh, Michelle DiGiacomo, Jane Phillips, and Patricia M. Davidson. 2007. Literature review: considerations in undertaking focus group research with culturally and linguistically diverse groups. *Journal of Clinical Nursing* 16, 6 (2007), 1000–1011. https://doi.org/10.1111/j.1365-2702.2006.01760.x

[90] Irene A. Harmsen, Liesbeth Mollema, Robert A. C. Ruiter, Theo G. W. Paulussen, Hester E. de Melker, and Gerjo Kok. 2013. Why parents refuse childhood vaccination: a qualitative study using online focus groups. *BMC Public Health* 13, 1 (2013), 8 pages. https://doi.org/10.1186/1471-2458-13-1183

[91] Sherif Hashem. 2019. Towards a National Cybersecurity Strategy : The Egyptian Case. *Journal of Systemics, Cybernetics and Informatics* 17, 3 (2019), 88–94. https://www.iiisci.org/journal/PDV/sci/pdfs/SA867CS19.pdf

[92] Sayed Hadi Hashemi, Mohammad Babaeizadeh, Mohsen Nowruzi, Hossein Hadian Jazi, Mohammad Shahmoradi, and Elaheh Biglar Beigi Samani. 2012. A Comprehensive Semi-automated Incident Handling Workflow. In *Proceedings of the 6th International Symposium on Telecommunications*. IEEE, 1065–1070. https://doi.org/10.1109/ISTEL.2012.6483144

[93] Otto Hellwig, Gerald Quirchmayr, Walter Hötzendorfer, Christof Tschohl, Edith Huber, Franz Vock, Florian Nentwich, Bettina Pospisil, Matthias Gusenbauer, and Gregor Langner. 2018. A GDPR Compliance Module for Supporting the Exchange of Information between CERTs. In *Proceedings of the 13th International Conference on Availability, Reliability and Security*. ACM, 7 pages. https://doi.org/10.1145/3230833.3233927

[94] Otto Hellwig, Gerald Quirchmayr, Edith Huber, Gernot Goluch, Franz Vock, and Bettina Pospisil. 2016. Major Challenges in Structuring and Institutionalizing CERT-communication. *Proceedings of the 2016 11th International Conference on Availability, Reliability and Security* 2, 661–667. https://doi.org/10.1109/ARES.2016.57

[95] Julian Higgins, James Thomas, Jacqueline Chandler, Miranda Cumpston, Tianjing Li, Matthew Page, and Vivian Welch. 2020. *Cochrane Handbook for Systematic Reviews of Interventions, Version 6.1*. Technical Report. https://training.cochrane.org/handbook/current

[96] Zakir Hossain, Golam Kibria Zaman, and Kazi Abu Taher. 2021. Cyber Emergency Response Team for Bangladesh. In *Proceedings of the 2021 International Conference on Information and Communication Technology for Sustainable Development*. IEEE, 477–480. https://doi.org/10.1109/ICICT4SD50815.2021.9396922

[97] Cathrine Hove, Marte Tarnes, Maria B. Line, and Karin Bernsmed. 2014. Information Security Incident Management: Identified Practice in Large Organizations. *Proceedings of the 8th International Conference on IT Security Incident Management & IT Forensics*, 27–46. https://doi.org/10.1109/IMF.2014.9

[98] Hsiu-Fang Hsieh and Sarah E. Shannon. 2005. Three Approaches to Qualitative Content Analysis. *Qualitative Health Research* 15, 9 (2005), 1277–1288. https://doi.org/10.1177/1049732305276687

[99] Jan Huck and Frank Breitinger. 2022. Wake Up Digital Forensics' Community and Help Combating Ransomware. *IEEE Security & Privacy* 20, 4 (2022), 61–70. https://doi.org/10.1109/MSEC.2021.3137018

[100] Sungsoo Hwang. 2008. Utilizing Qualitative Data Analysis Software: A Review of Atlas.ti. *Social Science Computer Review* 26, 4 (2008), 519–527. https://doi.org/10.1177/0894439307312485

[101] George Iakovakis, Constantinos Giovanni Xarhoulacos, Konstantinos Giovas, and Dimitris Gritzalis. 2021. Analysis and Classification of Mitigation Tools against Cyberattacks in COVID-19 Era. *Security and Communication Networks* 2021, Article 3187205 (2021), 21 pages. https://doi.org/10.1155/2021/3187205

[102] International Organization for Standardization (ISO). 2001. Software engineering – Product quality – Part 1: Quality model. web page. https://www.iso.org/standard/22749.html

[103] International Organization for Standardization (ISO). 2008. Software engineering – Software product Quality Requirements and Evaluation (SQuaRE) – Data quality model. web page. https://www.iso.org/standard/35736.html

[104] International Organization for Standardization (ISO). 2011. Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – System and software quality models. web page. https://www.iso.org/standard/35733.html

[105] International Organization for Standardization (ISO). 2018. Ergonomics of human-system interaction – Part 11: Usability: Definitions and concepts. web page. https://www.iso.org/standard/63500.html

[106] International Organization for Standardization (ISO). 2018. Information technology — Security techniques - Information security management systems — Overview and vocabulary. web page. https://www.iso.org/obp/ui/#iso:std:iso-iec:27000:ed-5:v1:en

[107] International Organization for Standardization (ISO). 2023. Information technology — Information security incident management — Part 1: Principles and process. web page. https://www.iso.org/obp/ui/#iso:std:iso-iec:27035:-1:ed-2:v1:en

[108] International Organization for Standardization (ISO). 2023. Information technology — Information security incident management — Part 2: Guidelines to plan and prepare for incident response. web page. https://www.iso.org/obp/ui/#iso:std:iso-iec:27035:-2:ed-2:v1:en

[109] International Telecommunication Union (ITU). [n.d.]. National CIRT. Web page. https://www.itu.int/en/ITU-D/Cybersecurity/Pages/national-CIRT.aspx

[110] International Telecommunication Unit. 2023. Global Cybersecurity Index 2020. web page. https://www.itu.int/epublications/publication/D-STR-GCI.01-2021-HTM-E

[111] Internet Governance Forum (IGF). 2014. Internet Governance Forum (IGF) 2014: Best Practice Forum on Establishing and Supporting Computer Security Incident Response Teams (CSIRT) for Internet Security. Online document. https://www.intgovforum.org/cms/documents/best-practice-forums/establishing-and-supporting-computer-emergency-response-teams-certs-for-internet-security/409-bpf-2014-outcome-document-computer-security-incident-response-teams/file

[112] Ron Iphofen and Martin Tolich. 2018. *The SAGE Handbook of Qualitative Research Ethics*. SAGE.

[113] ISO. 2022. Information security, cybersecurity and privacy protection — Information security controls. Web page. https://www.iso.org/standard/54533.html

[114] International Telecommunication Unit ITU. 2022. National CIRT. web page. https://www.itu.int/en/ITU-D/Cybersecurity/Pages/national-CIRT.aspx

[115] Martin Gilje Jaatun, Lars Bodsberg, Tor Olav Grøtan, and Marie Elisabeth Gaup Moe. 2020. An Empirical Study of CERT Capacity in the North Sea. In *Proceeding of the 2020 International Conference on Cyber Security and Protection of Digital Services*. IEEE, 8 pages. https://doi.org/10.1109/CyberSecurity49315.2020.9138865

[116] Islahuddin Jalal, Zarina Shukur, and Mohd Rosmadi Mokhtar. 2017. 3C-CSIRT Model: A Sustainable National CSIRT for Afghanistan. In *Proceedings of 6th International Conference on Electrical Engineering and Informatics (ICEEI), 2017*. IEEE, 1–4. https://doi.org/10.1109/ICEEI.2017.8312405

[117] Viktor Janevski. 2016. Shadowserver reports automated tool.

[118] Kamol Kaemarungsi, Nawattapon Yoskamtorn, Kitisak Jirawannakool, Nuttapong Sanglerdsinlapachai, and Chanin Luangingkasut. 2009. Botnet Statistical Analysis Tool for Limited Resource Computer Emergency Response Team. In *Proceedings of the 2009 Fifth International Conference on IT Security Incident Management and IT Forensics*. IEEE, 27–40. https://doi.org/10.1109/IMF.2009.13

[119] Renate M Kahlke. 2014. Generic qualitative approaches: Pitfalls and benefits of methodological mixology. *International journal of qualitative methods* 13, 1 (2014), 37–52.

[120] Panos Kampanakis. 2014. Security Automation and Threat Information-sharing Options. *IEEE Security & Privacy* 12, 5 (2014), 42–51. https://doi.org/10.1109/MSP.2014.99

[121] Hiroki Kashiwazaki. 2018. Personal Information Leak in a University, and Its Cleanup. In *Proceedings of the 2018 ACM SIGUCCS Annual Conference*. ACM, 43–50. https://doi.org/10.1145/3235715.3235727

[122] Manju Kaushik and Bhawana Mathur. 2014. Data Analysis of Students Marks with Descriptive Statistics. *International Journal on Recent and Innovation Trends in Computing and Communication* 2, 5 (2014), 1188–1190. https://doi.org/10.17762/ijritcc.v2i5.3136

[123] Kate Kelley, Belinda Clark, Vivienne Brown, and John Sitzia. 2003. Good practice in the conduct and reporting of survey research. *International Journal for Quality in Health Care* 15, 3 (2003), 261–266. https://doi.org/10.1093/intqhc/mzg031

[124] Naresh Khatri, Gordon D. Brown, and Lanis L. Hicks. 2009. From a Blame Culture to a Just Culture in Health Care. *Health Care Management Review* 34, 4 (2009), 312–322. https://doi.org/10.1097/HMR.0b013e3181a3b709

[125] Piotr Kijewski and Adam Kozakiewicz. 2011. Security Research at NASK: Supporting the Operational Needs of a CERT Team and More. In *Proceedings of 1st SysSec Workshop, 2011*. IEEE, 96–99. https://doi.org/10.1109/SysSec.2011.29

[126] Georgia Killcrece, Klaus-Peter Kossakowski, Robin Ruefle, and Mark Zajicek. 2003. *Organizational models for computer security incident response teams CSIRTs*. Handbook CMU/SEI-2003-HB-001. Carnegie Mellon University, Pittsburgh, Pennsylvania, USA. https://resources.sei.cmu.edu/asset_files/Handbook/2003_002_001_14099.pdf

[127] Sanghyun Kim and Gary Garrison. 2009. Investigating mobile wireless technology adoption: An extension of the technology acceptance model. *Information Systems Frontiers* 11 (2009), 323–333.

[128] Yujin Kim. 2011. The Pilot Study in Qualitative Inquiry: Identifying Issues and Learning Lessons for Culturally Competent Research. *Qualitative Social Work* 10, 2 (2011), 190–206. https://doi.org/10.1177/1473325010362001

[129] Barbara Kitchenham and Stuart Charters. 2007. *Guidelines for Performing Systematic Literature Reviews in Software Engineering.* Technical Report.

[130] Jenny Kitzinger. 1995. Qualitative Research: Introducing Focus Groups. *BMJ* 311, 7000 (1995), 299–302. https://doi.org/10.1136/bmj.311.7000.299

[131] Rick van der Kleij, Geert Kleinhuis, and Heather Young. 2017. Computer Security Incident Response Team Effectiveness: A Needs Assessment. *Frontiers in Psychology* (2017). https://doi.org/10.3389/fpsyg.2017.02179

[132] Herschel Knapp. 2017. *Intermediate Statistics using SPSS*. SAGE. https://doi.org/10.4135/9781071802625

[133] Erka Koivunen. 2010. "Why Wasn't I Notified?": Information Security Incident Reporting Demystified. In *Information Security Technology for Applications: 15th Nordic Conference on Secure IT Systems, NordSec 2010, Espoo, Finland, October 27-29, 2010, Revised Selected Papers (Lecture Notes in Computer Science, Vol. 7127).* Springer, 55–70. https://doi.org/10.1007/978-3-642-27937-9_5

[134] Farzan Kolini and Lech Janczewski. 2017. Clustering and Topic Modelling: A New Approach for Analysis of National Cyber Security Strategies. , 12 pages. https://aisel.aisnet.org/pacis2017/126/

[135] Toshihiro Komiyama, Shin'ichi Fukuzumi, Motoei Azuma, Hironori Washizaki, and Naohiko Tsuda. 2020. Usability of Software–Intensive Systems from Developers' Point of View. In *Human-Computer Interaction. Design and User Experience: Thematic Area, HCI 2020, Held as Part of the 22nd International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part I.* Springer, 450–463. https://doi.org/10.1007/978-3-030-49059-1_33

[136] Shunichi Konno. 2018. *Escalating Cyberattacks: Leveraging Threat Intelligence at NTT-CERT.* Technical Report. NTT CSIRT. 1–5 pages. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201805fa2_s.html

[137] Erik B. Korn, Douglas M. Fletcher, Erica M. Mitchell, Aryn A. Pyke, and Steven M. Whitham. 2021. Jack pandemus – Cyber Incident and Emergency Response During a Pandemic. *Information Security Journal: A Global Perspective* 30, 5 (2021), 294–307. https://doi.org/10.1080/19393555.2021.1980159

[138] Andrew G. Kotulic and Jan Guynes Clark. 2004. Why there aren't more Information Security Research Studies. *Information & Management* 41, 5 (2004), 597–607. https://doi.org/10.1016/j.im.2003.08.001

[139] Patrick Kral. 2012. *The Incident Handler's Handbook.* SANS White Paper. SANS Institute. https://www.sans.org/white-papers/33901/

[140] Marko Krstic, Milan Cabarkapa, and Aleksandar Jevremovic. 2019. Machine Learning Applications in Computer Emergency Response Team Operations. In *Proceedings of the 27th Telecommunications Forum*. IEEE, 4 pages. https://doi.org/10.1109/TELFOR48224.2019.8971040

[141] Richard A. Krueger. 2014. *Focus Groups: A Practical Guide for Applied Research.* SAGE.

[142] Marc Kührer, Christian Rossow, and Thorsten Holz. 2014. Paint it Black: Evaluating the Effectiveness of Malware Blacklists. In *Research in Attacks, Intrusions and Defenses: 17th International Symposium, RAID 2014, Gothenburg, Sweden, September 17-19, 2014, Proceedings*. Springer, 1–21. https://doi.org/10.1007/978-3-319-11379-1_1

[143] Seokcheol Lee and Taeshik Shon. 2017. Open Source Intelligence base Cyber Threat Inspection Framework for Critical Infrastructures. In *Proceedings of 2017 Future Technologies Conference*. IEEE, 1030–1033. https://doi.org/10.1109/FTC.2016.7821730

[144] Ifigeneia Lella, Eleni Tsekmezoglou, Rossen Svetozarov Naydenov, Cosmin Ciobanu, Apostolos Malatras, and Marianthi Theocharidou. 2022. *ENISA Threat Landscape Report 2022*. Technical Report. https://www.enisa.europa.eu/publications/enisa-threat-landscape-2022

[145] Theocharidou Marianthi Tsekmezoglou Eleni Malatras Apostolos Lella, Ifigeneia. 2022. ENISA Threat Landscape 2021. Web page. https://www.enisa.europa.eu/publications/enisa-threat-landscape-2021

[146] Anna Leppänen and Terhi Kankaanranta. 2020. Co-production of Cybersecurity: A Case of Reported Data System Break-ins. *Police Practice and Research* 21, 1 (2020), 78–94. https://doi.org/10.1080/15614263.2018.1525382

[147] R Barry Lewis. 2004. NVivo 2.0 and ATLAS.ti 5.0: A Comparative Review of Two Popular Qualitative Data-Analysis Programs. *Field Methods* 16, 4 (2004), 439–464. https://doi.org/10.1177/1525822X04269174

[148] Alessandro Liberati, Douglas G. Altman, Jennifer Tetzlaff, Cynthia Mulrow, Peter C. Gøtzsche, John P.A. Ioannidis, Mike Clarke, P J. Devereaux, Jos Kleijnen, and David Moher. 2009. The PRISMA Statement for Reporting Systematic Reviews and Meta-analyses of Studies that Evaluate Healthcare Interventions: Explanation and Elaboration. *BMJ* 339 (2009). https://doi.org/10.1136/bmj.b2700

[149] Maria B. Line, Inger Anne Tøndel, and Martin G. Jaatun. 2014. Information security incident management: Planning for failure. In *Proceedings of the 2014 8th International Conference on IT Security Incident Management & IT Forensics*. IEEE, 47–61. https://doi.org/10.1109/IMF.2014.10

[150] May O. Lwin, Jiahui Lu, Anita Sheldenkar, and Peter J. Schulz. 2018. Strategic Uses of Facebook in Zika Outbreak Communication: Implications for the Crisis and Emergency Risk Communication Model. *International Journal of Environmental Research and Public Health* 15, 9 (2018), 1974 pages. https://doi.org/10.3390/ijerph15091974

[151] Mingchao Ma, Kouril Daniel, Cyril L'Orphelin, Triantafyllidis Christos, and Michal Prochazka. 2012. EGI Security Monitoring. In *Proceedings of the International Symposium on Grids and Clouds (ISGC) 2012*. SISSA MEDIALAB SRL, 11 pages. https://pos.sissa.it/153/026/pdf

[152] Mark J Makowsky, Lisa M Guirguis, Christine A Hughes, Cheryl A Sadowski, and Nese Yuksel. 2013. Factors influencing pharmacists' adoption of prescribing: qualitative application of the diffusion of innovations theory. *Implementation Science* 8, 1 (2013), 1–11.

[153] Malaysian National Security Council. 2019. Malaysia Cyber Security Strategy 2020-2024. Governmental report. https://asset.mkn.gov.my/web/wp-content/uploads/sites/3/2019/08/MalaysiaCyberSecurityStrategy2020-2024Compressed.pdf

[154] Patrick Mana and Vasileios Friligkos. 2019. Eurocontrol/Eatm-Cert Services-Supporting Aviation To Better Manage Cyber Threats. In *Proceedings of Integrated Communications, Navigation and Surveillance Conference, 2019*. IEEE, 1–15. https://doi.org/10.1109/ICNSURV.2019.8735282

[155] Rima Masri and Monther Aldwairi. 2017. Automated malicious advertisement detection using VirusTotal, URLVoid, and TrendMicro. In *2017 8th International Conference on Information and Communication Systems (ICICS)*. 336–341. https://doi.org/10.1109/IACS.2017.7921994

[156] Jim A. McCall, Paul K. Richards, and Gene F. Walters. 1977. *Factors in Software Quality. Volume I. Concepts and Definitions of Software Quality.* Technical Report. General Electric Co. https://apps.dtic.mil/sti/citations/ADA049014

[157] Stefan Metzger, Wolfgang Hommel, and Helmut Reiser. 2011. Integrated Security Incident Management – Concepts and Real-world Experiences. In *Proceedings of 6th*

*International Conference on IT Security Incident Management and IT Forensics, 2011*. IEEE, 107–121. https://doi.org/10.1109/IMF.2011.15

[158] Clinton J Mielke and Hsinchun Chen. 2008. Botnets, and the cybercriminal underground. In *2008 IEEE International Conference on Intelligence and Security Informatics*. IEEE, 206–211.

[159] José P. Miguel, David Mauricio, and Glen Rodríguez. 2014. A Review of Software Quality Models for the Evaluation of Software Products. *International Journal of Software Engineering & Applications* 5, 6 (2014), 31–54. https://doi.org/10.5121/ijsea.2014.5603

[160] Matthew B. Miles and A. Michael Huberman. 1994. *Qualitative Data Analysis: An Expanded Sourcebook*. SAGE. https://uk.sagepub.com/en-gb/eur/qualitative-data-analysis/book246128

[161] Vikas Milhotra, Merrick S. Watchorn, Deepayan Chanda, et al. 2022. White Paper-Cybersecurity for Next-Generation Connectivity Systems–Rethinking Digital Architectures to Safeguard the Next Generation From Cybersecurity Breaches. (2022), 1–34. https://ieeexplore.ieee.org/document/9940224

[162] Jane Mills and Melanie Birks. 2014. Introducing Qualitative Research. *Qualitative Methodology: A Practical Guide* (2014), 3–16. https://doi.org/10.4135/9781473920163

[163] Sarandis Mitropoulos, Dimitrios Patsos, and Christos Douligeris. 2006. On Incident Handling and Response: A state-of-the-art approach. *Computers & Security* 25, 5 (2006), 351–370. https://doi.org/10.1016/j.cose.2005.09.006

[164] Sharifah Roziah Mohd Kassim, Sahrom Abu, and Amirah Mohd Omar. 2019. Measuring the Effectiveness of Phishing Detection Tool: Comparative Study on Pattern Matching and User Rating Technique. *Journal of Computers* 14, 4 (2019), 302–310. https://doi.org/10.17706/jcp.14.4.302-310

[165] Sharifah Roziah Mohd Kassim and Wira Zanoramy Ansiry Zakaria. 2014. Automating Big Data Analysis: Malaysia CERT Experience. In *Proceedings of the Tokyo International Conference on Engineering and Applied Sciences 2014*. https://doi.org/10.13140/2.1.3818.9125

[166] Sharifah Roziah Mohd Kassim, Wira Zanoramy A. Zakaria, and Nur Mohammad Kamil Mohammad Alta. 2016. Exploitation of Android Mobile Malware in Phishing Modus Operandi: A Malaysia Case Study. In *Proceedings of the 2nd International Conference on Electronics and Software Science*. SDIWC, 47–55.

https://www.academia.edu/29841520/Exploitation_of_Android_Mobile_Malware_in_Phishing_Modus_Operandi_A_Malaysia_Case_Study

[167] Sharifah Roziah Binti Mohd Kassim, Solahuddin Bin Shamsuddin, Shujun Li, and Budi Arief. 2022. How National CSIRTs Operate: Personal Observations and Opinions from MyCERT. In *Proceedings of the 2022 IEEE Conference on Dependable and Secure Computing*. IEEE, 2 pages. https://doi.org/10.1109/DSC54232.2022.9888803

[168] Sharifah Roziah Binti Mohd Kassim, Shujun Li, and Budi Arief. 2022. How National CSIRTs Leverage Public Data, OSINT and Free Tools in Operational Practices: An Empirical Study. *Cyber Security: A Peer-Reviewed Journal* 5, 3 (2022), 251–276. https://kar.kent.ac.uk/93768/

[169] Sharifah Roziah Binti Mohd Kassim, Shujun Li, and Budi Arief. 2022. Incident Response Practices Across National CSIRTs: Results from an Online Survey. *OIC-CERT Journal of Cyber Security* 4, 1 (2022), 63–80. https://www.oic-cert.org/en/journal/pdf/4/1/5.pdf

[170] Sami Mokaddem, Gerard Wagener, and Alexandre Dulaunoy. 2019. AIL - The design and implementation of an Analysis Information Leak Framework. In *Proceedings of 2018 IEEE International Conference on Big Data*. IEEE, 5049–5057. https://doi.org/10.1109/BigData.2018.8622074

[171] Roderick D. Mooi and Reinhardt A. Botha. 2016. A management model for building a computer security incident response capability. *SAIEE Africa Research Journal* 107, 2 (2016), 78–91. https://doi.org/10.23919/SAIEE.2016.8531544 (This paper is part of the proceedings of Information Security South African (ISSA) 2015.).

[172] Francesca Moretti, Liesbeth van Vliet, Jozien Bensing, Giuseppe Deledda, Mariangela Mazzi, Michela Rimondini, Christa Zimmermann, and Ian Fletcher. 2011. A standardized approach to qualitative content analysis of focus group discussions from different countries. *Patient Education and Counseling* 82, 3 (2011), 420–428. https://doi.org/10.1016/j.pec.2011.01.005

[173] Robert Morgus, Isabel Skierka, Mirko Hohmann, and Tim Maurer. 2015. *National CSIRTs and Their Role in Computer Security Incident Response*. Technical Report. https://newamerica.org/documents/1431/CSIRTs-incident-response.pdf

[174] Peter Murray and Kevin Donegan. 2003. Empirical linkages between firm competencies and organisational learning. *The Learning Organization* 10, 1 (2003), 51–62.

[175] Peter Murray-Rust. 2008. Open data in science. *Nature Precedings* (2008), 1–1. https://doi.org/10.1038/npre.2008.1526.1

[176] Michael D. Myers and Michael Newman. 2007. The Qualitative Interview in IS Research: Examining the Craft. *Information and Organization* 17, 1 (2007), 2–26. https://doi.org/10.1016/j.infoandorg.2006.11.001

[177] Radka Nacheva and Anita Jansone. 2020. Evaluation of Business Process Modelling Tools through Software Quality Metrics. *Baltic Journal of Modern Computing* 8, 4 (2020), 534–542. https://doi.org/10.22364/bjmc.2020.8.4.04

[178] Hidenori Nakai, Naohiko Tsuda, Kiyoshi Honda, Hironori Washizaki, and Yoshiaki Fukazawa. 2016. A SQuaRE-based Software Quality Evaluation Framework and its Case Study. In *Proceedings of the 2016 IEEE Region 10 Conference*. IEEE, 3704–3707. https://doi.org/10.1109/TENCON.2016.7848750

[179] Marcin Nawrocki, Maynard Koch, Thomas C Schmidt, and Matthias Wählisch. 2021. Transparent forwarders: an unnoticed component of the open DNS infrastructure. In *Proceedings of the 17th International Conference on emerging Networking EXperiments and Technologies*. 454–462.

[180] NCSC-UK. 2022. National Cyber Strategy 2022. Governmental report. https://www.gov.uk/government/publications/national-cyber-strategy-2022/national-cyber-security-strategy-2022

[181] Padmalata Nistala, Kesav Vithal Nori, and Raghu Reddy. 2019. Software Quality Models: A Systematic Mapping Study. In *Proceedings of the 2019 IEEE/ACM International Conference on Software and System Processes*. IEEE, 125–134. https://doi.org/10.1109/ICSSP.2019.00025

[182] Stephen Northcutt. 2001. *Computer Security Incident Handling: Step by Step, a Survival Guide for Computer Security Incident Handling*. Sans Institute.

[183] Mariam Nouh, Jason R.C. Nurse, Helena Webb, and Michael Goldsmith. 2019. Cybercrime investigators are users too! Understanding the socio-technical challenges faced by law enforcement. *Proceedings of the 2019 Workshop on Usable Security (USEC) at Network and Distributed System Security Symposium*, 11 pages. https://doi.org/10.14722/usec.2019.23032

[184] Monika Nowikowska. 2022. The Main Tasks of the Network of Computer Security Incident Response Teams in the Light of the Act on the National Cybersecurity System in Poland. In *Cybersecurity in Poland*. Springer, 223–241. https://doi.org/10.1007/978-3-030-78551-2_15

[185] Mathew Nyamagwa. 2010. A Layered Framework Approach to Mitigate Crimeware. (2010). https://commons.erau.edu/adfsl/2010/thursday/7/

[186] Megan Nyre-Yu, Robert S. Gutzwiller, and Barrett S. Caldwell. 2019. Observing Cyber Security Incident Response: Qualitative Themes from Field Research. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 63, 1 (2019), 437–441. https://doi.org/10.1177/1071181319631016

[187] Briony J. Oates. 2005. *Researching Information Systems and Computing.* SAGE. https://us.sagepub.com/en-us/nam/researching-information-systems-and-computing/book226898

[188] Michelle Odlum and Sunmoo Yoon. 2015. What Can we Learn About the Ebola Outbreak from Tweets? *American Journal of Infection Control* 43, 6 (2015), 563–571. https://doi.org/10.1016/j.ajic.2015.02.023

[189] Sean Oesch, Robert Bridges, Jared Smith, Justin Beaver, John Goodall, Kelly Huffer, Craig Miles, and Dan Scofield. 2020. An Assessment of the Usability of Machine Learning Based Tools for the Security Operations Center. In *Proceedings of the 2020 International Conferences on Internet of Things and IEEE Green Computing and Communications and IEEE Cyber, Physical and Social Computing and IEEE Smart Data and IEEE Congress on Cybermatics.* IEEE, 634–641. https://doi.org/10.1109/iThings-GreenCom-CPSCom-SmartData-Cybermatics50389.2020.00111

[190] (Forum of Incident Response and Inc.(FIRST) Security Teams. 2020. Traffic Light Protocol (TLP), Version 1.0. web page. https://www.first.org/tlp/

[191] Forum of Incident Response and Security TeamFIRST. [n.d.]. FIRST Members Around the World. web page. https://www.first.org/members/map

[192] City of London Police. 2022. NFIB Fraud and Cyber Crime Dashboard - 13 months of data. Web page. https://www.arcgis.com/apps/dashboards/0334150e430449cf8ac917e347897d46

[193] Kyohei Oguchi, Teru Yamazaki, and Masato Yamane. 2018. Incident Response Solution to Minimize Attack Damage. *NEC Technical Journal* 12, 2 (2018), 38–42. https://www.nec.com/en/global/techrep/journal/g17/n02/pdf/170208.pdf

[194] Paulo Oliveira, Fátima Rodrigues, and Pedro Rangel Henriques. 2005. A Formal Definition of Data Quality Problems. In *MIT International Conference on Information Quality.* http://mitiq.mit.edu/ICIQ/Documents/IQ%20Conference%202005/Papers/AFormalDefinitionofDQProblems.pdf

[195] Anthony J. Onwuegbuzie, Wendy B. Dickinson, Nancy L. Leech, and Annmarie G. Zoran. 2009. A Qualitative Framework for Collecting and Analyzing Data in Focus Group Research. *International Journal of Qualitative Methods* 8, 3 (2009), 1–21. https://doi.org/10.1177/160940690900800301

[196] Cliodhna O'Connor and Helene Joffe. 2020. Intercoder reliability in qualitative research: debates and practical guidelines. *International journal of qualitative methods* 19 (2020), 1609406919899220. https://doi.org/10.1177/1609406919899220

[197] Henrietta O'Connor and Clare Madge. 2017. Online Interviewing. In *The SAGE Handbook of Online Research Methods*. Vol. 2. SAGE, 416–434. https://doi.org/10.4135/9781473957992

[198] Harpreet Passi. 2018. Top 10 Popular Open Source Intelligence (OSINT) Tools. Web page. https://www.greycampus.com/blog/information-security/top-open-source-intelligence-tools

[199] Javier Pastor-Galindo, Pantaleone Nespoli, Felix Gómez Mármol, and Gregorio Martínez Pérez. 2020. The Not Yet Exploited Goldmine of OSINT: Opportunities, Open Challenges and Future Trends. *IEEE Access* 8 (2020), 10282–10304. https://doi.org/10.1109/ACCESS.2020.2965257

[200] Celeste Lyn Paul. 2014. Human-Centered Study of a Network Operations Center: Experience Report and Lessons Learned. In *Proceedings of the 2014 ACM Workshop on Security Information Workers*. ACM, 39–42. https://doi.org/10.1145/2663887.2663899

[201] Paweł Pawlinski and Andrew Kompanek. 2016. Evaluating Threat Intelligence Feeds. Slides presented at the 2016 FIRST Technical Colloquium for Threat Intelligence. https://www.first.org/resources/papers/munich2016/kompanek-pawlinski-evaluating-threat-ntelligence-feeds.pdf

[202] Geoff Payne and Judy Payne. 2004. Coding Qualitative Data. In *Key Concepts in Social Research*. SAGE. https://doi.org/10.4135/9781849209397

[203] Geoff Payne and Judy Payne. 2004. Content analysis. In *Key Concepts Social Research*. SAGE, 51–55. https://doi.org/10.4135/9781849209397

[204] Geoff Payne and Judy Payne. 2004. Quantitative Methods. In *Key Concepts in Social Research*. SAGE, 181–186. https://doi.org/10.4135/9781849209397

[205] William H. Percy, Kim Kostere, and Sandra Kostere. 2015. Generic Qualitative Research in Psychology. *The Qualitative Report* 20, 2 (2015), 76–85. https://doi.org/10.46743/2160-3715/2015.2097

[206] Alexander Peychev. 2022. What is the Measured Response to a Cyber Attack on Critical Infrastructures? web page. https://www.nonproliferation.eu/wp-content/uploads/2022/05/Essay-Al.-Peychev.pdf

[207] Zane Pokorny (Ed.). 2019. *The Threat Intelligence Handbook: Moving Toward a Security Intelligence Program* (2nd ed.). CyberEdge Group, LLC. https://go.recordedfuture.com/book-2

[208] Raminta Pranckutė. 2021. Web of Science (WoS) and Scopus: The Titans of Bibliographic Information in Today's Academic World. *Publications* 9, 1, Article 12 (2021). https://doi.org/10.3390/publications9010012

[209] PricewaterhouseCoopers. 2022. *Cyber Threats 2022: A Year in Retrospect.* Technical Report. PricewaterhouseCoopers. https://www.pwc.com/gx/en/issues/cybersecurity/cyber-threat-intelligence/cyber-year-in-retrospect.html

[210] Darren Quick and Kim Kwang Raymond Choo. 2018. Digital Forensic Intelligence: Data Subsets and Open Source Intelligence (DFINT+OSINT): A timely and Cohesive mix. *Future Generation Computer Systems* 78 (2018), 558–567. https://doi.org/10.1016/j.future.2016.12.032

[211] Reischaga, Charles Lim, and Yohanes Syailendra Kotualubun. 2020. Uncovering Malware Traits Using Hybrid Analysis. In *Proceedings of the 2021 International Conference on Engineering and Information Technology for Sustainable Industry.* 1–6.

[212] Francsico Reyes, Walter Fuertes, Freddy Tapia, Theofilos Toulkeridis, Hernán Aules, and Ernesto Pérez. 2018. A BI Solution to Identify Vulnerabilities and Detect Real-Time Cyber-Attacks for an Academic CSIRT. In *Intelligent Computing: Proceedings of the Computing Conference, Volume 2, 2018.* Springer, 1135–1153. https://doi.org/10.1007/978-3-030-01177-2_82

[213] Roger J. Rezabek. 2000. Online Focus Groups: Electronic Discussions for Research. In *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, Vol. 1. https://doi.org/10.17169/fqs-1.1.1128

[214] Thea Riebe, Marc-André Kaufhold, and Christian Reuter. 2021. The Impact of Organizational Structure and Technology Use on Collaborative Practices in Computer Emergency Response Teams: An Empirical Study. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2, Article 478 (2021), 30 pages. https://doi.org/10.1145/3479865

[215] Robert Roohparvar. 2020. Use of OSINT tools for security and their functions. Web page. http://www.infoguardsecurity.com/use-of-osint-tools-for-security-and-their-functions/

[216] Robin Ruefle, Audrey Dorofee, David Mundie, Allen D. Householder, Michael Murray, and Samuel J. Perl. 2014. Computer Security Incident Response Team Development and Evolution. *IEEE Security & Privacy* 12, 5 (2014), 16–26. https://doi.org/10.1109/MSP.2014.89

[217] Erin Ruel, William Edward Wagner III, and Brian Joseph Gillespie. 2016. *The Practice of Survey Research: Theory and Applications*. SAGE. https://doi.org/10.4135/9781483391700

[218] Johnny Saldaña. 2015. *The Coding Manual for Qualitative Researchers*. SAGE.

[219] Janet Salmons. 2014. *Qualitative Online Interviews: Strategies, Design, and Skills*. SAGE.

[220] Jo Samanta and Ash Samanta. 2018. *A Focus Group Content Analysis Study Exploring Cultural and Faith Based Values at End of Life*. SAGE. https://doi.org/10.4135/9781526439239

[221] Helen Sampson. 2004. Navigating the waves: the usefulness of a pilot in qualitative research. *Qualitative Research* 4, 3 (2004), 383–402. https://doi.org/10.1177/1468794104047236

[222] Willem E. Saris and Irmtraud N. Gallhofer (Eds.). 2014. *Design, Evaluation, and Analysis of Questionnaires for Survey Research* (2nd ed.). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118634646

[223] Benjamin Saunders, Julius Sim, Tom Kingstone, Shula Baker, Jackie Waterfield, Bernadette Bartlam, Heather Burroughs, and Clare Jinks. 2018. Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality & quantity* 52 (2018), 1893–1907.

[224] Bruce Schneier. 2014. The Future of Incident Response. *IEEE Security & Privacy* 12, 5 (2014), 96–96. https://doi.org/10.1109/MSP.2014.102

[225] Thomas Schreck. 2018. *IT Security Incident Response: Current State, Emerging Problems, and New Approaches*. Friedrich-Alexander-Universitaet Erlangen-Nuernberg (Germany).

[226] Andreas Sfakianakis, Christos Douligeris, Louis Marinos, Marco Lourenço, and Omid Raghimi. 2019. *ENISA Threat Landscape Report 2018: 15 Top Cyberthreats and Trends*. Technical Report. https://doi.org/10.2824/622757

[227] Shadowserver Foundation. [n.d.]. Shadowserver. website. https://www.shadowserver.org/

[228] Nivedita Shinde and Priti Kulkarni. 2021. Cyber Incident Response and Planning: A Flexible Approach. *Computer Fraud & Security* 2021, 1 (2021), 14–19. https://doi.org/10.1016/S1361-3723(21)00009-9

[229] Brijendra Singh and Suresh Prasad Kannojia. 2013. A Review on Software Quality Models. In *Proceedings of the 2013 International Conference on Communication Systems and Network Technologies*. IEEE, 801–806. https://doi.org/10.1109/CSNT.2013.171

[230] Vandana Singh and Alexander Thurman. 2019. How Many Ways Can we Define Online Learning? A Systematic Literature Review of Definitions of Online Learning (1988-2018). *American Journal of Distance Education* 33, 4 (2019), 289–306. https://doi.org/10.1080/08923647.2019.1663082

[231] Florian Skopik, Giuseppe Settanni, and Roman Fiedler. 2016. A Problem Shared is a Problem Halved: A survey on the Dimensions of Collective Cyber Defense through Security Information Sharing. *Computers & Security* 60 (2016), 154–176. https://doi.org/10.1016/j.cose.2016.04.003

[232] Software Engineering Institute, Carnegie Mellon University. [n.d.]. National Computer Security Incident Response Teams (CSIRTs). web page. https://www.sei.cmu.edu/our-work/cybersecurity-center-development/index.cfm

[233] Software Engineering Institute, Carnegie Mellon University. 2023. CERT Division. webpage. https://www.sei.cmu.edu/about/divisions/cert/

[234] L. Karen Soiferman. 2010. Compare and Contrast Inductive and Deductive Research Approaches. web page. https://eric.ed.gov/?id=ED542066

[235] Eugene H. Spafford. 1989. The Internet worm program: An analysis. *ACM SIGCOMM Computer Communication Review* 19, 1 (1989), 17–57. https://doi.org/10.1145/66093.66095

[236] Jonathan M. Spring and Phyllis Illari. 2021. Review of Human Decision-making during Computer Security Incident Analysis. *Digital Threats: Research and Practice* 2, 2, Article 11 (2021), 47 pages. https://doi.org/10.1145/3427787

[237] ISO Standard. [n.d.]. ISO/IEC 27002:2013 Information technology — Security techniques — Code of practice for information security controls. Web page. https://www.iso.org/standard/54533.html

[238] Statista. 2023. Monetary Loss of Companies in the United States as a Result of Cyber Attacks as of 2022. web page. `https://www.statista.com/statistics/1334399/us-common-results-of-cyber-attacks/`

[239] Julie Steinke, Balca Bolunmez, Laura Fletcher, Vicki Wang, Alan J. Tomassetti, Kristin M. Repchick, Stephen J. Zaccaro, Reeshad S. Dalal, and Lois E. Tetrick. 2015. Improving Cybersecurity Incident Response Team Effectiveness using Teams-based Research. *IEEE Security & Privacy* 13, 4 (2015), 20–29. `https://doi.org/10.1109/MSP.2015.71`

[240] David W. Stewart and Prem Shamdasani. 2017. Online Focus Groups. *Journal of Advertising* 46, 1 (2017), 48–60. `https://doi.org/10.1080/00913367.2016.1252288`

[241] Kate Stewart and Matthew Williams. 2005. Researching Online Populations: The Use of Online Focus Groups for Social Research. *Qualitative Research* 5, 4 (2005), 395–416. `https://doi.org/10.1177/1468794105056916`

[242] Borce Stojkovski, Gabriele Lenzini, Vincent Koenig, and Salvador Rivas. 2021. What's in a Cyber Threat Intelligence Sharing Platform? A Mixed-methods User Experience Investigation of MISP. In *Proceedings of the 2021 Annual Computer Security Applications Conference*. ACM, 385–398. `https://doi.org/10.1145/3485832.3488030`

[243] Manoj Wadhwa Suman. 2014. A Comparative Study of Software Quality Models. *International Journal of Computer Science and Information Technologies* 5, 4 (2014), 5634–5638. `https://ijcsit.com/docs/Volume%205/vol5issue04/ijcsit20140504177.pdf`

[244] Sathya Chandran Sundaramurthy, John McHugh, Xinming Simon Ou, Siva Rajagopalan, and Michael Wesch. 2014. An Anthropological Approach to Studying CSIRTs. *IEEE Security & Privacy* 12, 5 (2014), 52–60. `https://doi.org/10.1109/MSP.2014.84`

[245] Rey LeClerc Sveinsson. 2022. Top 10 Data Breaches So Far in 2022. Web page. `https://ermprotect.com/blog/top-10-data-breaches-so-far-in-2022/`

[246] Muhammad Tariq, Baber Aslam, Imran Rashid, and Adeela Waqar. 2013. Cyber Threats and Incident Response Capability- A Case Study of Pakistan. In *Proceedings of 2nd National Conference on Information Assurance, 2013*. IEEE, 15–20. `https://doi.org/10.1109/NCIA.2013.6725319`

[247] Jaak Tepandi, Mihkel Lauk, Janar Linros, Priit Raspel, Gunnar Piho, Ingrid Pappel, and Dirk Draheim. 2017. *The Data Quality Framework for the Estonian Public Sector and Its Evaluation.* Springer Berlin Heidelberg, 1–26. https://doi.org/10.1007/978-3-662-56121-8_1

[248] Renata Tesch. 2013. *Qualitative Research: Analysis Types and Software.* Routledge. https://doi.org/10.4324/9781315067339

[249] Jason E. Thomas. 2019. A Case Study Analysis of the Equifax Data Breach. (12 2019), 12 pages. https://doi.org/10.13140/RG.2.2.16468.76161

[250] Inger Anne Tøndel, Maria B. Line, and Martin Gilje Jaatun. 2014. Information Security Incident Management: Current Practice as Reported in the Literature. *Computers & Security* 45 (2014), 42–57. https://doi.org/10.1016/j.cose.2014.05.003

[251] Shohei Toyama and Masayuki Hirayama. 2018. User Interface Design Method Considering UI Device in Internet of Things System. In *Proceedings of the 2018 6th International Conference on Future Internet of Things and Cloud Workshops.* IEEE, 1–6. https://doi.org/10.1109/W-FiCloud.2018.00007

[252] Roumen Trifonov, Radoslav Yoshinov, Slavcho Manolov, Georgi Tsochev, and Galya Pavlova. 2019. Artificial Intelligence Methods Suitable for Incident Handling Automation. In *MATEC Web of Conferences*, Vol. 292. EDP Sciences, 01044. https://doi.org/10.1051/matecconf/201929201044

[253] Naohiko Tsuda, Hironori Washizaki, Kiyoshi Honda, Hidenori Nakai, Yoshiaki Fukazawa, Motoei Azuma, Toshihiro Komiyama, Tadashi Nakano, Hirotsugu Suzuki, Sumie Morita, Katsue Kojima, and Akiyoshi Hando. 2019. WSQF: Comprehensive Software Quality Evaluation Framework and Benchmark Based on SQuaRE. In *Proceedings of the 2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice.* IEEE, 312–321. https://doi.org/10.1109/ICSE-SEIP.2019.00045

[254] Liam Tung. 2022. Microsoft: Russia has Launched Hundreds of Cyberattacks against Ukraine. Web page. https://www.zdnet.com/article/microsoft-russia-has-launched-hundreds-of-cyberattacks-against-ukraine/

[255] Kota Uehara, Kohei Mukaiyama, Masahiro Fujita, Hiroki Nishikawa, Takumi Yamamoto, Kiyoto Kawauchi, and Masakatsu Nishigaki. 2019. Basic Study on Targeted E-mail Attack Method Using OSINT. In *Advanced Information Networking and Applications: Proceedings of the 33rd International Conference on Advanced Information Networking and Applications (AINA-2019) (Advances in*

*Intelligent Systems and Computing, Vol. 926).* Springer, 1329–1341. https://doi.org/10.1007/978-3-030-15032-7_111

[256] University of Kent. 2022. Research Ethics and Governance. Webpage. https://www.kent.ac.uk/research/research-ethics-and-governance

[257] US-CERT. [n.d.]. Traffic Light Protocol (TLP) Definitions and Usage. Web page. https://www.cisa.gov/tlp

[258] Mojtaba Vaismoradi and Sherrill Snelgrove. 2019. Theme in Qualitative Content Analysis and Thematic Analysis. In *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, Vol. 20. DEU. https://nordopen.nord.no/nord-xmlui/bitstream/handle/11250/2627867/Vaismoradi.pdf?sequence=4

[259] Paul Valladares, Walter Fuertes, Freddy Tapia, Theofilos Toulkeridis, and Ernesto Perez. 2017. Dimensional Data Model for Early Alerts of Malicious Activities in a CSIRT. In *2017 International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS)*, Vol. 49. IEEE, 1–8. https://doi.org/10.23919/SPECTS.2017.8046771

[260] Martijn Van der Heide. 2017. Establishing a CSIRT. In *Forum of Incident Response and Security Teams (FIRST).*

[261] Rick Van der Kleij, Geert Kleinhuis, and Heather Young. 2017. Computer Security Incident Response Team Effectiveness: A Needs Assessment. *Frontiers in Psychology* 8, 2179 (2017), 8 pages. https://doi.org/10.3389/fpsyg.2017.02179

[262] Michel van Eeten, Qasim Lone, Hadi Asghari TUD, and Hadi Asghari. [n.d.]. CERT. ([n. d.]).

[263] Javier Verdugo and Moisés Rodríguez. 2020. Assessing Data Cybersecurity using ISO/IEC 25012. *Software Quality Journal* 28, 3 (2020), 965–985. https://doi.org/10.1007/s11219-019-09494-x

[264] Verizon. 2022. *2022 Data Breach Investigations Report.* Technical report. Verizon. https://www.verizon.com/business/resources/reports/dbir/

[265] Antonio Vetrò, Lorenzo Canova, Marco Torchiano, Camilo Orozco Minotas, Raimondo Iemma, and Federico Morando. 2016. Open Data Quality Measurement Framework: Definition and Application to Open Government Data. *Government Information Quarterly* 33, 2 (2016), 325–337. https://doi.org/10.1016/j.giq.2016.02.001

[266] Clarke Victoria, Braun Virginia, and Hayfield Nikki. 2017. Thematic Analysis. *Journal of Positive Psychology* 12, 3 (2017), 297–98. https://doi.org/10.1080/17439760.2016.1262613

[267] William Villegas-Ch., Ivan Ortiz-Garces, and Santiago Sánchez-Viteri. 2021. Proposal for an Implementation Guide for a Computer Security Incident Response Team on a University Campus. *Computers* 10, 8, Article 102 (2021), 23 pages. https://doi.org/10.3390/computers10080102

[268] Martijn Visser, Nees Jan van Eck, and Ludo Waltman. 2021. Large-scale Comparison of Bibliographic Data Sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic. *Quantitative Science Studies* 2, 1 (2021), 20–41. https://doi.org/10.1162/qss_a_00112

[269] Václav Vostrovskỳ, Jan Tyrychtr, and Roman Kvasnička. 2020. Open Data Quality Management Based on ISO/IEC SQuaRE Series Standards in Intelligent Systems. In *Applied Informatics and Cybernetics in Intelligent Systems: Proceedings of the 9th Computer Science On-line Conference 2020, Volume 3*. Springer, 625–631. https://doi.org/10.1007/978-3-030-51974-2_58

[270] Fathul Wahid. 2007. Using the technology adoption model to analyze internet adoption and use among men and women in Indonesia. *The Electronic Journal of Information Systems in Developing Countries* 32, 1 (2007), 1–8.

[271] Y. M. Wara and D. Singh. 2015. A Guide to Establishing Computer Security Incident Response Team (CSIRT) for National Research and Education Network (NREN). *African Journal of Computing & ICT* 8, 2 (2015), 1–8. https://afrjcict.net/wp-content/uploads/2017/08/vol-8-no-2-june-2015ppi-198.pdf

[272] William A. Ward Jr and Buvaneswari Venkataraman. 1999. Some Observations on Software Quality. In *Proceedings of the 37th Annual Southeast Regional Conference*. ACM, 2–9. https://doi.org/10.1145/306363.306367

[273] Douglas Wells. 2016. Taking Stock of Subjective Narratives Surrounding Modern OSINT. In *Open Source Intelligence Investigation: From Strategy to Implementation*. Springer, 57–65. https://doi.org/10.1007/978-3-319-47671-1_5

[274] Rodrigo Werlinger, Kasia Muldner, Kirstie Hawkey, and Konstantin Beznosov. 2010. Preparation, Detection, and Analysis: the Diagnostic Work of IT Security Incident Response. *Information Management & Computer Security* 18, 1 (2010), 26–42. https://doi.org/10.1108/09685221011035241

[275] Moira J. West-Brown, Don Stikvoort, Klaus-Peter Kossakowski, Georgia Killcrece, Robin Ruefle, and Mark Zajicek. 2003. *Handbook for Computer Security Incident Response teams (CSIRTs)* (2nd ed.). Technical Report CMU/SEI-2003-HB-002. Software Engineering Institute, Carnegie Mellon University, Pittsburg, USA. https://resources.sei.cmu.edu/asset_files/Handbook/2003_002_001_14102.pdf

[276] Marilyn Domas White and Emily E. Marsh. 2006. Content analysis: A flexible methodology. *Library Trends* 55, 1 (2006), 22–45. https://doi.org/10.1353/lib.2006.0053

[277] White House. 2018. National Cyber Strategy of the United States of America. Governmental report. https://digital.library.unt.edu/ark:/67531/metadc1259394/

[278] World Health Organisation WHO. 2023. Coronavirus Diseses (COVID-19) Pandemic. Web page. https://www.who.int/europe/emergencies/situations/covid-19

[279] Johannes Wiik, Jose J. Gonzalez, and Klaus Peter Kossakowski. 2006. Effectiveness of Proactive CSIRT Services. In *Proceedings of the 18th Annual FIRST Conference*. FIRST, 13 pages. https://www.first.org/conference/2006/papers/kossakowski-klaus-papers.pdf

[280] Claes Wohlin. 2014. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*. 1–10. https://doi.org/10.1145/2601248.2601268

[281] Muhammd Mudassar Yamin and Basel Katt. 2019. A survey of automated information exchange mechanisms among CERTs. In *Proceedings of the 5th Collaborative European Research Conference (CERC 2019)*, Vol. 2348. CEUR Workshop Proceedings, 311–322. http://hdl.handle.net/11250/2624127

[282] Jacob Young and Sahar Farshadkhah. 2021. Backdoors & Breaches: Using a Tabletop Exercise Game to Teach Cybersecurity Incident Response. In *Proceedings of the EDSIG Conference ISSN*, Vol. 2473. 4901. https://proc.iscap.info/2021/pdf/5562.pdf

[283] Tangxiao Yuan, Kondo Hloindo Adjallah, Alexandre Sava, Huifen Wang, and Linyan Liu. 2021. Issues of Intelligent Data Acquisition and Quality for Manufacturing Decision-Support in an Industry 4.0 Context. In *Proceedings of the*

*2021 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, Vol. 2. IEEE, 1200–1205. https://doi.org/10.1109/IDAACS53288.2021.9660957

[284] Zahri Yunos, Ramona Susanty Ab Hamid, and Mustaffa Ahmad. 2016. Development of a Cyber Security Awareness Strategy Using Focus Group Discussion. In *Proceedings of the 2016 SAI Computing Conference*. IEEE, 1063–1067. https://doi.org/10.1109/SAI.2016.7556109

[285] Mohammad Zarour. 2020. A Rigorous User Needs Experience Evaluation Method based on Software Quality Standards. *Telkomnika Journal* 18, 5 (2020). https://doi.org/10.12928/TELKOMNIKA.v18i5.16061

[286] Junwen Zhu and Weishu Liu. 2020. A Tale of Two Databases: The Use of Web of Science and Scopus in Academic Papers. *Scientometrics* 123, 1 (2020), 321–335. https://doi.org/10.1007/s11192-020-03387-8

[287] Shuofei Zhu, Ziyi Zhang, Limin Yang, Linhai Song, and Gang Wang. 2020. Benchmarking label dynamics of virustotal engines. In *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*. 2081–2083.

[288] Chaim Zins. 2007. Conceptual Approaches for Defining Data, Information, and Knowledge. *Journal of the American Society for Information Science and Technology* 58, 4 (2007), 479–493. https://doi.org/10.1002/asi.20508

[289] Hugo Zylberberg and Alexander Klimburg. 2015. Cyber Security Capacity Building: Developing Access. (2015).

# Appendix A

# Evaluation Result of Tools - Product Quality and Quality in Use

## Evaluation of VirusTotal and Hybrid Analysis Tools – Product Quality

| **Criteria** | **Metrics** | **Result and Value** | |
|---|---|---|---|
| **Security** | | **VirusTotal** | **Hybrid Analysis** |
| Confidentiality | Not Applicable (NA) because end users of this service (national CSIRTs) are not concerned about the confidentiality of the data to authorised users only. | NA | NA |
| Integrity | Applicable because end users care about data integrity (such as data is not sniffed, modified or manipulated by a  third party).<br>Metric 1: If the data transmission is encrypted end to end using HTTPS (Binary: Yes and No)<br>Metric 2: If the online service is hacked by a malicious party who can manipulate the data before it is sent to the end users<br>(Binary: Yes and No) | <br>Yes<br>No | <br>Yes<br>No |
| Authenticity | Applicable because end users care about the authenticity of the server.<br>Metric 1: If the HTTPS protocol is used to provide service authentication<br>(Binary: Yes and No)<br>Metric 2: To what extent can end users be sure about the real identity of the service provider or the developer if there is no service provider (e.g., a company's registration record, paid membership official contract documents, natural person's real-world identity), whose value can be one of the following:<br>• the real identity can be fully obtained<br>• the real identity can be partially obtained (e.g., just an online account, but not real-world identity)<br>• the real identity cannot be obtained | Yes. HTTPS is used<br>'The real identity can be fully obtained' from a DomainTools whois search, with the below result:<br>IP: 74.125.34.46<br>IP Location: California – Mountain View – Google<br>ASN: AS15169, Google, US (registered Mar 30, 2000)<br>Company information provided on its website as:<br>Virus Total, USA | Yes. HTTPS is used<br>'The real identity can be fully obtained' from a DomainTools whois search, with the below result:<br>IP: 104.18.34.183<br>IP Location: California – San Jose – CloudFlare Inc.<br>ASN: AS13335, Cloudflare-net, US (registered Jul 14, 2010)<br>Company information provided on its website as:<br>Hybrid Analysis, Germany |
| **Usability** | | | |

| | | | |
|---|---|---|---|
| Learnability | Applicable because end users care about how easily they can learn about the features of the service.<br>Metric: The amount of time a typical end user (such as a member of technical staff at a national CSIRT) takes to learn the task of submitting a file to the tool in order to get the result or output after submitting the file.<br>Measurement method: this metric can be measured by asking several typical end users about the time taken to learn the task of submitting an URL to the tool in order to get the result or output after submitting the URL. of the tool. | ‘5 minutes’ (judged based on the researcher’s testing of learning the task of submitting a file because trying the UI is straightforward for any technical person) | ‘8 minutes’ (judged based on the researcher’s testing of learning the task of submitting a file because trying the UI is straightforward for any technical person) |
| Operability | Applicable because end users want to use a system with easy control -- User Interface (UI)<br>Metric 1: If the service has a GUI that can be used by a junior-technical person (Binary: Yes and No)<br>Metric 2: If the service has an online interface that can be used by anyone with Internet access (Binary: Yes and No)<br>Metric 3: If the service has an API so that a member of the technical staff at a national CSIRT can automate the use of the service (Binary: Yes and No) | Yes<br>Yes<br>Yes | Yes<br>Yes<br>Yes |
| User Interface Aesthetics | Applicable as technical staff of national CSIRTs are concerned about the simplicity of the menu to make the task less complicated.<br>Metric: How the result is presented Value: ‘In a table’, ‘Colour codes legend’, ‘Buttons for detail result’ | Result is presented in ‘structured table’ form on the page. Has ‘colour code’, e.g. Red – indicates the file submitted is detected as malware, Green – indicates not malware. Colour code of Confidence Level of the result: Red – indicates low confidence level, | Result is ‘not presented in table form’. Has ‘colour code’, e.g. Red – indicates the file submitted is detected as malware, Green – indicates Clean. Has a ‘button to display details’ of the result |

| | | | |
|---|---|---|---|
| | | Green – indicates high confidence level. Has a ‘button to display details’ of the result | |
| Accessibility | Applicable as technical staff of national CSIRTs is concerned if they can have access to the tool for investigation purposes.<br>Metric: If access to the tool is enabled for users with cognitive or physical impairments<br>Value: ‘Font size’, ‘Sound control’, ‘Colour palettes’, ‘Brightness of the web page can be adjusted’, ‘Font size of the web page can be enlarged’ | ‘Has option for night and day screen viewing’, ‘Font size of the web page can be enlarged’ | ‘Font size of the web page can be enlarged’ |
| **Maintainability** | | | |
| Maintainability | Applicable since the service can be updated and the provider can make the updates visible to users.<br>Metric: If the tool shows the current year of the tool, visible to users<br>Value: ‘Current year 2003 is shown’, ‘Current year 20023 is not shown’ | ‘Current year 2023 is not shown’ on the web page | ‘Current year is shown’ as: ©2023 Hybrid Analysis |
| Supportability | Applicable as users are concerned that they can get support and help when any issues arise while using the tool.<br>Metric: What type of support is provided by the tool<br>Value: ‘24x7 Live Support’, ‘Online Support’, ‘Chat Bot’ | ‘24x7 Live Support’, ‘Online Support’, ‘Chat Bot’ | ‘Online Support’ |
| Analysability | This is applicable as users should be able to know causes of failures that might occur while using the tool.<br>Metric: If causes of failures can be identified<br>(Binary: Yes and No) | Yes | Yes |
| Modifiability | Not applicable to free tools | | |
| **Compatibility** | | | |
| Interoperability | Applicable as users is concerned that the tool can be integrated with other third-party applications. | | |

| | | | |
|---|---|---|---|
| | Metric: If the tool allows integration and information exchange with third-party applications<br>Value: ‘Has an API for integration with third-party applications’, ‘Has an export feature that allows data exchange’ | ‘Has API integration’ with 14 third party applications | ‘Has API integration’ with 1 third party applications |
| **Functionality** | | | |
| Functionality | Applicable as users are concerned that the tool functions cover all the specified tasks in the tool and meet the user's objective.<br>Metric: if the tool has a file scanning feature and performs the specified functionality (feature) accordingly<br>(Binary: Yes and No) | Yes. The tools is able to scan a file and diagnose a malicious file as malicious<br>In contrast, when input a clean file, the file is detected as clean | Yes. The tools is able to scan a file and diagnose a malicious file as malicious<br>In contrast, when input a clean file, the file is detected as clean |
| Performance Efficiency | | | |
| Time behaviour | Applicable because end users care how fast they can get the results.<br>Metric: The amount of time taken to obtain the result when input a file or URL (from the start to the return of the results) | ‘2 seconds’ (according to researcher’s testing) | ‘9 seconds’ (according to researcher’s testing) |
| Capacity | Applicable because end users care about the size of the file that can be uploaded to the website to perform the analysis.<br>Metric: ‘maximum file size that can be uploaded for analysis’<br>Measurement method: from the official documents and testing it yourself | 650 MB can be uploaded to the website | 100 MB to the website |
| Money | Not applicable (NA) because this is a free service. | NA | NA |
| Human Effort | Not applicable (NA) because this is an automated online tool. | NA | NA |
| Material | Applicable as users want to know how many inputs or materials need | ‘1 file input’ is sufficient to generate the result’ | ‘1 file input’ is sufficient to generate the result’ |

| | | | |
|---|---|---|---|
| | to be provided to get the desired result.<br>Value: ‘1 file input’, ‘More than 1 file input’ | | |
| Reliability | | | |
| Reliability | Applicable as users are concerned that the tool provides accurate result. | | |
| | Metric: The tool returns accurate result for the input given.<br>(Binary: Yes and No) | Yes. | Yes. |
| Availability | Applicable as users are concerned the service is available to perform the tasks whenever required. | | |
| | Metric: If the tool can run if any third-party libraries or data sources it relies upon are down.<br>(Binary: Yes and No) | Yes.  The tool can run and gives the required result | Yes. The tool can run and gives the required result |
| | Metric: If the tool can run offline even if its web server is down<br>Value: Not able to evaluate this as there was no circumstances of the web server was down during the evaluation exercise. By description of the tool, the value should be ‘No’, as the tool is an online service (not offline) | No | No |
| | Metric: The availability of the tool to perform its intended function without failure during the evaluation period.<br>Value: ‘99.999% uptime’ | ‘99.999% uptime’ during the evaluation period | ‘99.999% uptime’ during the evaluation period |
| Compliance | Applicable as the tool needs to comply with any standards or industry practices. | | |
| | Metric: If the tool complies with one or more specific industry standards or practices.<br>Value: ‘Name of Standard’, ‘Not found’ | ‘Not found’ | ‘Not found’ |
| Certification | Applicable, as a certified or accredited tool is always a good sign of its quality. | ‘Not found’ | ‘Not found’ |

| | | | |
|---|---|---|---|
| | Metric: If it is certified and publicised, what certifying body and Standard.<br>Value: ‘Name of Certification’, ‘Not found’ | | |

## Evaluation of VirusTotal and Hybrid Analysis Tools – Quality in Use

| **Criteria** | **Metrics** | **Result and Values** | |
|---|---|---|---|
| **Context Coverage** | | VirusTotal | Hybrid Analysis |
| Flexibility | Applicable as users are concerned about the tool’s flexibility to specific users (e.g. non-expert users) to achieve intended goals.<br>Metrics: If users perceive the software is flexible for non-expert users.<br>(Binary: Yes and No)<br>These perceptions can be obtained from ``External reports" and ``Estimations with Co-workers from a national CSIRT".<br>If the above external reports or estimation with co-workers could not be obtained, the value is ``Not found". However, if a national CSIRT wants to evaluate a tool based on this criterion, it is best to survey and gather opinions and insights on the software's flexibility to get the value. | Yes | No |
| **Usability** | | | |
| Satisfaction | | | |
| User Experience | Applicable as users are concerned that they perceive and are satisfied with the usefulness of the result.<br>Metric: If users perceive the result from the tool as useful.<br>(Binary: Yes and No)<br>Measurement: Based on how users perceive the usefulness. These perceptions can be obtained from ``External reports" and ``Estimations with Co-workers from a national CSIRT".<br>If the above external reports or estimation with co-workers could not be obtained, the | Yes | Yes |

| | | | |
|---|---|---|---|
| | value is ``Not found". However, if a national CSIRT wants to evaluate a tool based on this criterion, it is best to survey to get the value. | | |
| Usefulness | Applicable as users are concerned that they perceive and are satisfied with the usefulness of the result.<br>Metric: If users perceive the result from the tool as useful.<br>(Binary: Yes and No)<br>Measurement: Based on how users perceive the usefulness. These perceptions can be obtained from ``External reports" and ``Estimations with Co-workers from a national CSIRT".<br>If the above external reports or estimation with co-workers could not be obtained, the value is ``Not found". However, if a national CSIRT wants to evaluate a tool based on this criterion, it is best to survey to get the value. | Yes | Yes |
| Trust | Applicable as users care if they have confidence in the tool that it functions as it is supposed to be.<br>Metric: if users have confidence that the tool will function as it should.<br>(Binary: Yes and No)<br>Measurement: Based on users' trust and confidence. This trust and confidence level can be obtained from ``External reports" and ``Estimations with Co-workers from a national CSIRT".<br>If the above external reports or estimation with co-workers could not be obtained, then the value is ``Not found". However, if a national CSIRT wants to evaluate a tool based on this criterion, it is best to survey to get the value. | Yes | Yes |
| Comfort | Not applicable (NA) for online services. | NA | NA |
| **Effectiveness** | | | |
| Effectiveness | Applicable as users care if the tool meets the specified requirements and performs the functions.<br>Metric: if the tool fulfils the specified requirements and performs the specified functions correctly.<br>(Binary: Yes and No) | Yes | Yes |
| **Freedom from Risk** | | | |

|  |  |  |  |
|---|---|---|---|
| Sustainability | Applicable as users want to make sure the tool is sustained without depletion (with the very basic feature) by the Publisher.<br>Metric: how the sustainability of the tool is guaranteed. | Maintained by a large, well-known organisation | Dependent on single company's existence and business-driven decision |
| Harm from use | Applicable as users want to make sure no negative consequences regarding health, safety, finances or the environment that result from the use of the system.<br>Metric: if the tool posed health hazards to the user.<br>(Binary: Yes and No) | No | No |
| **Popularity** | Applicable as users care if the tool is used by others, such as national CSIRTs or other security organisations.<br>Metrics: How many security organisations or national CSIRTs use the tool.<br>Value: ‘Large’, ‘Medium’, ‘Low’<br>This information can be obtained from ‘External reports’ and ‘Informal discussions with the security community’.<br>If the above external reports or discussions could not be obtained, the value is ‘Not known’. However, if a national CSIRT wants to evaluate a tool based on this criterion, it is best to survey to get the value. | Used by ‘large’ number of national CSIRTs (based on the researcher’s informal discussion with national CSIRTs) | Used by ‘medium’ number of national CSIRTs (based on the researcher’s informal discussion with national CSIRTs) |

# Appendix B

# Evaluation Result of Data Source

**Evaluation of Shadowserver Data Source**

| Criteria | Metrics and Values | Result and Value |
|---|---|---|
| | | Shadowserver Data Feed |
| Confidentiality | Not applicable (NA) to end users of this service (national CSIRTs) because they don't have concerns about the confidentiality of the data feeds (which are public) | NA |
| Accuracy | Applicable as users care that the data has accurate information about an incident such as an indicator of compromise (IOC).<br>Metric: If the data accurately indicates or tells about an incident<br>Value: ‘Presence of malware URL’, ‘Presence of phishing URL’, ‘Hash value of malware’ | ‘Presence of malware URL’, ‘Presence of phishing URL’, ‘Hash value of malware’ |
| Precision | Applicable as users care that the data is precise enough for taking further action, such as takedowns.<br>Metric: If data has specific details about an incident<br>Value: ‘Has complete URL’, ‘Has source IP address with date and timestamp’ (to detect hosts using DHCP IP address), ‘Has a hash value of malware’, ‘Has Internet protocol type’, ‘Has a port number’, ‘Has IP geo-location’ | ‘Has complete URL’, ‘Has source IP address with date and time stamp’, ‘Has a hash value of malware’, ‘Has Internet protocol type’, ‘Has a port number’, ‘Has IP geolocation’ |
| Understandability | Applicable to ensure the data is in a format that makes data easily understood and facilitate analysis.<br>Metric: The format that can be understood by human users to facilitate analysis<br>Value: ‘Can be exported as a CSV file’, ‘Displayed structured in a table’ | ‘Can be exported as a CSV file’, ‘Displayed structured in a table’ |
| Currentness | Applicable as users are concerned that the data is current and not outdated.<br>Metric: Shows the current date and timestamp of the data<br>Value: ‘Shows current date’ | ‘Shows current date - 12/02/2023 1:28:00 AM’ which is the date of evaluation |
| Completeness | Applicable as users care about the sufficiency of data for further incident response action.<br>Metric: Data is sufficient for a further incident response such as escalation to Service Provider<br>Value: ‘Has a complete URL’, ‘Has IP address, ‘Has date’, ‘Has timestamp’, ‘Has ASN number’, ‘Name of malware’, ‘Has malware hash value’ | ‘Has a complete URL’, ‘Has IP address’, ‘Has date’, ‘Has timestamp’, ‘Has ASN number’, ‘Name of malware’, ‘Has malware hash value’ |

|  |  |  |
|---|---|---|
| Credibility | Applicable as users care the data is credible to support incident response.<br>Metric: Source of data is from a known reliable data provider.<br>Value: ‘Yes/No’<br>Metric: The data has information it intends to provide.<br>Value: ‘Yes/No’ | ‘Yes’<br>‘Yes’ |
| Efficiency | Applicable as users care that the data is efficient in terms of less time consuming to identify indicators to initiate incident response.<br>Metric: Time taken to identify an indicator of compromise from the data.<br>Value: ‘Number - second/minute/hour’ | ‘6 seconds’ |