Ethical Issues in Cybersecurity: Employing red teams, responding to ransomware attacks and attempting botnet takedowns.

Gwenyth Morgan MSc.

A dissertation submitted in fulfilment of the requirements for the award of

Doctor of Philosophy (Ph.D.)

To the

Dublin City University

School of Theology, Philosophy & Music

Supervisors:

Dr. Bert Gordijn

Institute of Ethics, Dublin City University

Dr. Dave Lewis

School of Computer Science, Trinity College Dublin

Dr. Joss Moorkens

School of Applied Language and Intercultural Studies, Dublin City University

September 2021

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of PhD is entirely my own work, and that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed:

ID No.: 17210200

Date: 2 September 2021

Acknowledgements

Preparing this dissertation has been one of the most challenging and intellectually stimulating experiences of my life. I will be forever indebted to my primary supervisor Bert Gordijn who presented me with this incredible opportunity. Without Bert's support and endless patience this thesis would not have come to completion. I would like to thank Bert for introducing me to the esteemed members of the CANVAS project and having confidence in me to lead and contribute to the CANVAS project deliverables. To everyone on the CANVAS project, I am very grateful for the immeasurable influence you had on me in the earlier days of my research, particularly those who were involved in preparing the Reference Curriculum. I am not sure that we were aware of the enormity of the task at hand, but we got there in the end!

I would like to thank my two secondary supervisors, Joss Moorkens and Dave Lewis for helping me in my pursuit to bridge the divide between cybersecurity and business ethics. Saying that I was fortunate to have these two wise gentlemen as my supervisors is an understatement as they never failed to share their infinite wisdom upon reviewing draft document after draft document. Bert, Joss and Dave connected me to the Institute of Ethics at DCU and to the Ethics and Privacy Working Group at the ADAPT Centre. Both networks enabled me to learn from many brilliant minds namely Fiachra O'Brolchain and Declan O'Sullivan. Fiachra acted as mentor, confidante and philosophical interpreter. While Declan shared his enthusiasm in those earlier days of my research when I presented this huge idea with no clue how to tackle it.

It was an absolute pleasure working with Renaat Verbruggen who kindly helped me tease through the intricacies of cybersecurity attacks. I must also say a big thank you to Alexey Kirichenko, Erka Koivunen, Petros Efstathopoulos and Florian Grunow who are fighting the real fight against cybercriminals in prominent industry roles. I am very grateful for the many crucial conversations we had which helped me narrow the focus on ethical issues in cybersecurity in respect of cybersecurity attacks, ethical hacking and ubiquitous computing.

Without the unconditional love and support that I received from family and friends, this thesis would never have reached completion: Daniela, my astute friend who encouraged me to start this journey all those years ago. Every day since, Daniela has consistently been a source of light and positivity. My dear friend Jenny with whom I embarked on many peaceful walks that often ended in memorable nights out and pep-talks. My inordinately intelligent, witty younger brother and friend Evan who was often forced to indulge in every ethical angle that popped into my head. My oldest brother Gareth whom I greatly admire as a human being. Gareth's unwavering support and loyalty kept me moving forward. Lynette, Adam and Jake brought vibrancy and sheer joy to my life during our family video calls during the COVID-

19 pandemic while Eoin kept reminding me to believe in myself and to always think about the bigger picture. Sean, the most reliable and grounding partner in crime. We country hopped back and forth from New Zealand to the Netherlands to Ireland and back again over the last four years while I worked on the CANVAS project and wrote this dissertation. Somehow, Sean still managed to pick up the pieces whenever I needed him to and for that I will always be grateful. And finally, my parents, who have given me so much in life. Mam and Dad provided me with the care, education, critical outlook, and perseverance I needed to complete this thesis. Even though we were not residing in the same country or continent throughout my thesis, they guided my every step.

Table of Contents

ACKNOWLEDGEMENTS ... III

ABBREVIATIONS ... VII

ABSTRACT ... IX

CHAPTER 1 INTRODUCTION ... 1

1.1 RESEARCH QUESTIONS ... 1

1.2 METHODOLOGY ... 2

1.3 RELEVANCE ... 4

1.4 OUTLINE ... 7

CHAPTER 2 ETHICAL ISSUES IN CYBERSECURITY IN THE BUSINESS DOMAIN ... 8

2.1 INTRODUCTION ... 8

2.2 METHODOLOGY ... 8

2.3 RESULTS ... 11

2.3.1 CYBERSECURITY THREATS TO BUSINESSES ... 12
2.3.2 WHAT ARE THE ETHICAL ISSUES IN CYBERSECURITY IN THE BUSINESS DOMAIN? ... 14

2.4 DISCUSSION ... 23

2.4.1 CONFLICTS OF INTERESTS ... 23
2.4.2 BLIND SPOTS ... 29

2.5 CONCLUSION ... 30

CHAPTER 3 ETHICAL THEORY ... 32

3.1 INTRODUCTION ... 32

3.2 STAKEHOLDER THEORY & PRINCIPLE OF FAIR PLAY ... 32

3.2.1 ANALYSING RAWLS' PRINCIPLE OF FAIR PLAY ... 33
3.2.2 PHILLIPS' PRINCIPLE OF STAKEHOLDER FAIRNESS ... 35
3.2.3 THE COOPERATIVE SCHEME, STAKEHOLDER STATUS & STAKEHOLDER OBLIGATIONS ... 35

3.3 NORMATIVE, DERIVATIVE AND NON-STAKEHOLDERS ... 36

3.3.1 NORMATIVE STAKEHOLDERS ... 36
3.3.2 DERIVATIVE STAKEHOLDERS ... 37
3.3.3 NON-STAKEHOLDERS ... 37

3.4 PRIORITIZING STAKEHOLDERS ... 38

CHAPTER 4 EMPLOYING A RED TEAM ... 40

4.1 INTRODUCTION ... 40

4.2 RED TEAMING ... 40

4.2.1 ETHICAL HACKING ... 40
4.2.2 ADVANCED SECURITY ... 41
4.2.3 RED TEAMS VERSUS PEN TESTERS ... 43
4.2.4 HOW RED TEAMING WORKS ... 44

4.3 ETHICAL ANALYSIS ... 46

4.3.1 NORMATIVE STAKEHOLDERS' INTERESTS ... 47
4.3.2 DERIVATIVE STAKEHOLDERS' INTERESTS ... 58
4.3.3 PRIORITISING CONFLICTING STAKEHOLDERS' INTERESTS ... 60

4.4 CONCLUSION ... 62

CHAPTER 5 RESPONDING TO RANSOMWARE ATTACKS 63

5.1 INTRODUCTION 63

5.2 RANSOMWARE 64

5.2.1 A GROWING THREAT & A CRIMINAL OFFENCE 65
5.2.2 IMPACT & COST 66
5.2.3 WHO IS A TARGET? 68
5.2.4 HOW ARE VICTIMS RESPONDING? 69

5.3 ETHICAL ANALYSIS 70

5.3.1 NORMATIVE STAKEHOLDER INTERESTS 71
5.3.2 IS THERE A DIFFERENCE BETWEEN PAYING AND NEGOTIATING? 75
5.3.3 DERIVATIVE STAKEHOLDER INTERESTS 77
5.3.4 PRIORITISING CONFLICTING STAKEHOLDER INTERESTS 79

5.4 CONCLUSION 81

CHAPTER 6 ATTEMPTING A BOTNET TAKEDOWN 82

6.1 INTRODUCTION 82

6.2 DDOS ATTACKS 84

6.2.1 A GROWING THREAT TO ORGANISATIONS 84
6.2.2 DOS, DDOS, BOTNETS & TAKEDOWNS 88
6.2.3 DEFENCE METHODS: FROM PASSIVE TO ACTIVE 90

6.3 ETHICAL ANALYSIS 95

6.3.1 NORMATIVE STAKEHOLDERS' INTERESTS 95
6.3.2 DERIVATIVE STAKEHOLDERS INTERESTS 107
6.3.3 PRIORITISATION OF CONFLICTING STAKEHOLDERS' INTERESTS 112

6.4 CONCLUSION 112

CHAPTER 7 CONCLUSION 116

7.1 SUMMARY OF THE RESULTS 116

7.2 LIMITATIONS 120

7.3 FUTURE RESEARCH 120

BIBLIOGRAPHY 122

Abbreviations

ACD	Active Cyber Defence
ACM	The Association for Computing Machinery
ACSC	The Australian Cyber Security Centre
API	Application Programming Interfaces
APTs	Advanced Persistent Threats
ARPANET	The Advanced Research Projects Agency Network
CaaS	Crime as a Service
CANVAS	Constructing an Alliance for Value-driven Cybersecurity
CDN	Content Delivery Network
CERT-EU	European Cybersecurity Emergency Response Team
CISO	Chief Information Security Officer
CSP	Cloud Service Provider
C&C	Command and Control
DD4BC	Distributed Denial of Service for Bitcoin
DDoS	Distributed Denial of Service
DHB	District Health Board
DJF	Dirt Jumper Family
DNS	Domain Name System
DoS	Denial of Service
GDP	Gross Domestic Product
GDPR	The General Data Protection Regulation
HSE	Health Service Executive
ICT	Information Communication Technology
IDPM	The British Computer Society and the Institute of Data Processing Management
IEEE	The Institute of Electrical and Electronic Engineers
IFIP	The International Federation of Information processors
IoCs	Indicators of Compromise
IP	Internet Protocol
IPS	Intrusion Prevention Systems
IS	Information Security
ISP	Internet Service Provider
IT	Information Technology
Mbps	Megabytes per second
MIT	Massachusetts Institute of Technology
MNCs	Multinational Corporations
NHS	National Health Service
OS	Operating System
OSSTMM	Open Source Security Testing Methodology Manual
PEN	Penetration
PETs	Privacy Enhancing Technologies
PII	Personal Identifiable Information
PoSF	Principle of Stakeholder Fairness
PR	Public Relations
P2P	Peer-to-Peer
RDDoS	Ransomware Distributed Denial of Service
RSA	Rivest-Shamir-Ardleman
SMEs	Small-to-Medium-sized Enterprises
ST	Stakeholder Theory
Tbps	Terabytes per second
TCP	Transmission Control Protocol
TTPs	Tactics, Techniques and Procedures

UID	User Identification
UK	United Kingdom
USA	The United States of America
USD	United States Dollar
VMPs	Vulnerability Management Programs
VRPs	Vulnerability Reward Programs
WoS	Web of Science

Abstract

Gwenyth Morgan. Ethical issues in cybersecurity: Employing red teams, responding to ransomware attacks & attempting botnet takedowns.

The following four research questions are analysed in this thesis: What are the ethical issues that arise in cybersecurity in the business domain? Is it ethically appropriate for organisations to employ red teams to find security vulnerabilities? What is the ethically appropriate organisational response to a ransomware attack? Is it ethically appropriate for organisations to attempt a botnet takedown in response to a DDoS attack? The first research question is answered by way of a literature review which reveals that many ethical issues arise in cybersecurity in the business domain. The second, third and fourth research questions are analysed using a strategic method described by Robert A Phillips. This method, based on stakeholder theory and the political theory of John Rawls, provides a philosophical basis for stakeholder legitimacy and the prioritisation of stakeholders' interests should conflict of interests amongst stakeholders arise. This method can be replicated by decision-makers to determine ethically appropriate courses of action to take.

Chapter 1 Introduction

1.1 Research Questions

The four research questions analysed in this thesis are as follows:

1) What are the ethical issues that arise in cybersecurity in the business domain? Are there any blind spots in the ethical literature that are worthy of further ethical deliberation?
2) Is it ethically appropriate for organisations to employ red teams to find security vulnerabilities?
3) What is the ethically appropriate organisational response to a ransomware attack?
4) Is it ethically appropriate for organisations to attempt a botnet takedown in response to a DDoS attack?

The first research question is two-fold: 1) what are the ethical issues that arise in cybersecurity in the business domain and 2) are there any blind spots in the ethical literature that are worthy of further ethical deliberation? The remaining three research questions relate to three blind spots identified in the ethical literature: ethical hacking (employing red teams), cybersecurity threats (responding to ransomware and Distributed Denial of Service (DDoS) attacks) and ubiquitous devices (DDoS attacks are enabled by the widespread use of ubiquitous devices i.e., devices connected to the internet also known as the Internet of Things (IoTs)).

The first blind spot raises the issue of organisations[1] employing ethical hackers known as red teams to test the two main security vulnerabilities in organisations: people and technology. Red teaming is an "authorised, adversary-based assessment for defensive purposes" (Sandia National Laboratory, 2011). By identifying vulnerabilities, red teams can provide insight and guidance on how to improve the security posture of the organisation. To identify vulnerabilities red teams often employ deceptive and manipulative tactics that target people and technology in the sponsor organisation. The aim of the ethical analysis in Chapter 4 is to determine whether employing red teams is ethically appropriate.

The second blind spot concerns the ethics of responding to a specific cybersecurity threat called ransomware. Ransomware attacks[2] are a growing threat to organisations, and at the time of writing organisations are left to decide for themselves how they should respond to a ransomware attack. Some

[1] The term "organisation" is interchangeably used with business, institution, company or firm unless otherwise specified as Small-to-Medium-sized Enterprises (SME) or Multinational Corporations (MNCs). It is a general term which will be used to encapsulate all privately and publicly owned organisations that reside within the business sector.

[2] Attack refers broadly to operations in cyberspace that attempt to compromise or impair the confidentiality, availability, or integrity of electronic information, information systems, services, or networks (Hoffman & Levite, 2017).

choose to pay, while others decide to not pay or negotiate a lower ransom[3]. The aim of examining this cybersecurity threat in Chapter 5 is to examine the ethical issues that arise from organisations choosing to pay, not pay or negotiate a ransom with cybercriminals and determine the ethically appropriate response to a ransomware attack.

The third blind spot relates to the growing threat of DDoS attacks to organisations. This threat emanates from widespread use of insecure devices connected to the internet, the IoTs. Organisations can actively respond to DDoS attacks by attempting to takedown the source from which the attack is staged, a process known as a botnet takedown. There is no systematic business ethics analysis of attempting a botnet takedown and how it may affect key stakeholders' interests. The aim of ethically analysing an attempted botnet takedown in Chapter 6 is to ascertain whether it is ethically appropriate.

1.2 Methodology

A two-pronged methodological approach is adopted.

1) To answer the first research question, the author of this thesis completed a literature review titled the Ethics of Cybersecurity in Business. This review is published in the white paper, "Ethics and Cybersecurity" (Yaghmaei, et al., 2017). The findings from this published[4] review can be found in Chapter 2 of this thesis and the methodological details of the review are described in Section 2.2.

2) Ethical theory

   Robert Phillips' Stakeholder Theory (ST) is used to answer the second, third and fourth research questions (see Chapters 4, 5 & 6). The details of Phillips' Stakeholder Theory are described in Chapter 3. It seems apt to ethically analyse research questions two, three and four from a business ethics perspective, one that places stakeholder interests at the centre of the decision-making process. ST is a business ethics theory which suggests that in addition to shareholders there is a multiplicity of groups who have a stake in the operation of a business – all of whom merit

[3] At the time of submission (June 2021), new stories permeated the globe surrounding a ransomware that successfully targeted the Health Service Executive (HSE) in Ireland (Halpin & Humphries, 2021; Mehta, 2021; MacNamee, 2021; Perlroth, 2021). The HSE attack forced the HSE to shut down all IT systems to reduce collateral damage caused by the attack (Halpin & Humphries, 2021). In the absence of IT, healthcare workers are being forced to use paper records to keep services operational which is causing severe disruption to services during the COVID-19 pandemic (Perlroth, 2021). €16.3million was demanded by the attackers in return for access and the HSE publicly stated they will not pay the ransom (MacNamee, 2021).

[4] While writing of this thesis, the author also published 2 chapters included in the *White Paper,– Attitudes and Opinions Regarding Cybersecurity in Health, Busines and Critical Infrastructure* (Wenger & et al, 2017); and a book chapter titled, 'A Care-Based Stakeholder Approach to Ethics of Cybersecurity in Business', in Christen, M., Gordijn, B. & Loi M. (Eds.), *The Ethics of Cybersecurity*, Springer, New York, 2020, 119-138 (Christen, et al., 2020).

consideration in managerial decision making i.e., shareholders, employees, customers, suppliers, and the local community (Freeman et al., 2011). It is understood that the decision-making authority in an organisation may be management, a senior leadership team or otherwise. For the purposes of simplicity, these organisational representatives are referred to as management or the organisation throughout this thesis.

Ed Freeman, credited as the founding father of ST, argues that all organisations must endeavour to create value for stakeholders by promoting their interests and putting those interests at the centre of business decisions (Freeman, 1984). This view is similar to the main tenet of utilitarianism which weighs an action's potential effects on those affected by the action. The action that results in the greatest amount of good for everyone is considered the ethical action. In contrast to utilitarianism, Robert Phillips amalgamates Ed Freeman's concept of ST with the moral and political theory of John Rawls (Rawls, 1964). Phillips' research makes it abundantly clear that some stakeholders merit more consideration than others based on their contribution to the firm and whether they are a co-operator in a mutually beneficially scheme (Phillips, 1997). Phillips' notion of stakeholder theory thus migrates away from utilitarianism as it is not for the benefit of the majority, but for the benefit of particular stakeholders with whom the organisation has a particular relationship. The particulars of such a relationship are outlined in Phillips' work on fair play and stakeholder legitimacy (see section 3.3 for more details) which is based on Rawls' principle of fair play (Phillips, 1997) – a concept which was first mentioned by John Stuart Mill in 1859 (Mill, 1989).

Without rehashing information that is comprehensively described in Chapter 3, it is helpful at this juncture to mention that Philips provides us with the tools we need to demarcate between those to whom the organisation has a moral obligation to consider in the decision-making process, a group he calls normative stakeholders, and those to whom the organisation has a derived obligation, referred to as derivative stakeholders (Phillips, 2003a) . This prescription not only enables us to determine to whom obligations are owed but allows management to prioritise stakeholders' interests based on equitable contribution to the firm i.e., those who contribute the most should be in receipt of the larger share or voice. If it is not possible for management to measure contribution, Philips explains that managers should take action that advances normative stakeholders' interests, supports the continuation of the cooperative scheme and is likely to achieve the assent of all normative stakeholders (see Chapter 3 for more details).

1.3 Relevance

Organisations have greatly benefited from adopting new and innovative Information Technology (IT). Benefits extend to increasing productivity, reducing operational costs, improving customer service, and maximising overall revenue. However, the risk of staying connected broadens the spectrum through which organisations can be targeted by criminals. Criminals can target servers, computers, connected devices and the people that have access to these technologies. Due to cybersecurity attacks being one of the biggest threats to organisational growth and the global economy (Ernst Young, 2019), establishing and maintaining secure information systems, protecting the data held within those systems and protecting the people who have access to those systems (a practice commonly referred to as cybersecurity) has become a strategic priority for organisations.

Cybersecurity is not only a relevant topic in the realm of business, but it is a relevant topic of discussion in any sector that relies upon technology to execute daily functions and operations. What is concerning for all organisations is that cybersecurity attacks are rising in frequency and severity; the cost of cybersecurity attacks is increasing, the public disclosure of a cybersecurity incident impacts several key stakeholders, and the two main cybersecurity weaknesses in organisations are people and technology with no silver bullet for managing either.

1) The frequency and severity of attacks is increasing.

   Cybersecurity attacks usually involve a malicious attempt to break into a system, interrupt service or steal data. There is a growing list of tools that can be used to execute an attack including exploit kits, malware[5] such as ransomware or a DDoS attack. DDoS attacks are increasing in frequency and severity. Nexusguard reported a 278% increase in DDoS attacks in the second quarter of 2020 when compared to the same period in 2019. This was a 542% increase compared to the previous quarter (Nexusguard, 2020).

2) The cost of cybersecurity attacks is increasing.

   The cost and collateral damage caused by cybersecurity attacks can be detrimental to a business, depending on the scale of the attack, the organisation targeted and how far the virus has spread. Financial losses are increasing for businesses who have suffered a cybersecurity attack. For example, the average loss from a cybersecurity attack in 2018 was 1.23 million USD which increased to 1.41 million USD in 2019 (Kaspersky, 2019a). Fast forward to 2021, the average cost of a ransomware payment in the first quarter of 2021 was USD 220,298 (Coveware, 2021) and the total

[5] Malware is a broad term that is used to describe various types of malicious programs. Malicious programs can perform undesirable operations such as compromising a computer or stealing data. Common types of malware include trojans, viruses, worms and spyware (ENISA, 2021c).

cost of a ransomware attack (including device and network cost, lost opportunity and ransom paid) is averaging at USD 1.85million (Sophos, 2021).

3) Disclosing cybersecurity incidents impacts several key stakeholders.

   Disclosing cybersecurity incidents can negatively affect market share and stock price. This means that disclosing incidents may affect stakeholders' i.e., shareholders and employees' financial interest in the success of the firm. Organisations may attempt to hide a cybersecurity incident with the intention of avoiding the deleterious effects associated with the public disclosure of an incident. One example of underreporting an incident occurred in 2016 when Uber tried to cover up news that 57 million driver and rider accounts were breached. Information pertaining to the breach was not disclosed until a year after the incident took place (Perlroth & Isaac, 2018). The Uber breach occurred in the United States of America (USA), where all fifty states have enacted legislation[6] that requires private and governmental entities to notify individuals of security breaches of Personal Identifiable Information (PII) (Greenburg, 2018). This case suggests that although legislation compels organisations to divulge breach information, it does not guarantee compliance. It also suggests that many stakeholders can be affected by disclosure or a cover up, including customers whose PII was leaked.

4) The main cybersecurity weaknesses are people and technology.

   Organisations typically rely upon people and technology for growth and prosperity. These assets are also the main weaknesses in cybersecurity (F-Secure, 2018). According to a 2018 PwC report, employees are responsible for 27% of all cybersecurity incidents (PwC(UK), 2018). This figure includes malicious[7] and non-malicious[8] employee behaviour, both of which are collectively referred to as the insider threat. Cybercriminals use social engineering techniques to manipulate employees to gain access to data or devices. According to the Information Security Institute, the most popular social engineering attacks are executed via email, social media, instant messaging, and SMS to trick victims into providing sensitive information or visiting malicious URLs to compromise their systems (Infosec, 2019). A good example of a social engineering attack is a ransomware attack. Ransomware attacks can be launched in various ways, but one popular method used to launch an attack is to disguise the ransomware virus as an email attachment. All it takes for the infection to start is for an unwitting employee to download and open the attachment – it is that easy. A ransomware attack

[6] It is worth mentioning that Europe might see similar cases to Uber despite the enactment of the General Data Protection Regulation ("GDPR"). The GDPR came into effect in Europe in May 2018 which compels businesses to share breach information with a supervisory authority within 72 hours of the breach occurring and failure to do so will result in a fine of "up to 4% of the total worldwide annual turnover" (European Parliament, 2016).

[7] A malicious employee intends to cause harm to an organisation.

[8] A non-malicious employee can cause unintentional harm through their lack of cybersecurity knowhow or through a lack of applying cybersecurity knowhow.

can steal valuable trade information, delete data, or encrypt data and devices which can cause significant disruption to an organisation which is evidenced by the Irish HSE attack in 2021 (see Section 1.1 for more details). Kaspersky's 2019 survey on industrial companies shows that of the 282 participants surveyed, 70% consider ransomware as their greatest concern, trumping targeted attacks like Advanced Persistent Threats (APTs), sabotage and threats from supply chain or partners such as third parties (Kaspersky, 2019c).

Cybercriminals target technology by exploiting technical vulnerabilities to gain access, disrupt operations or cause damage to an organisation. Criminals often exploit internet-based vulnerabilities to launch DDoS attacks. DDoS attacks are used to disrupt internet activity. This means a successful DDoS attack can cause a website to slow or shut down completely which can create havoc for internet facing organisations i.e., gaming, gambling, streaming services. DDoS attacks are often executed in conjunction with other attack vectors such as ransomware for DDoS (RDDoS). The problem is that DDoS attacks continue to grow in frequency and size, for which there is currently no quick fix.

5) There is no silver bullet.

To combat the insider threat some organisations are choosing to secretly monitor their employees (ENISA, 2016). The problem with secret surveillance is it can be perceived by employees as an infringement of privacy and autonomy in the workplace (Brey, 2007). In respect of technical vulnerabilities, MNCs host competitions known as either Vulnerability Reward Programs (VRPs) or "bug bounties" (Tripwire, 2015). These are structured gaming competitions and are often invite-only events for highly adept security researchers where large monetary rewards (ranging from thousands to hundreds of thousands of USD) are offered to those who can find technical vulnerabilities in the host's infrastructure. However, hosting such an event can be costly and is limited to financially flexible MNCs. SMEs, who represent 99% of all businesses in Europe do not have the financial flexibility to host such events (European Commission, 2018). In fact, SMEs may not have the resources to employ or even reward financially motivated security researchers, never mind host a large bug bounty program. Security capabilities between SMEs and MNCs will widen such that MNCs continue to fortify defences while SMEs become more vulnerable to attacks.

The specific research questions analysed in this thesis are relevant for the following reasons:

1) The first research question highlights the many ethical issues that arise in cybersecurity in the business domain and areas that require further ethical scrutiny (Chapter 2). This contributes to the limited ethical research in the areas of cybersecurity and business.
2) The second research question focuses on organisations choice to employ ethical hackers such as red teams to test the two main vulnerabilities in organisations, people and technology (Chapter 4).

The stakeholder analysis addresses the limited ethical literature on the ethics of employing red teams whilst providing insight into the way in which red teams engage with organisations and how such engagements may impact the interests of various stakeholders.

3) The third and fourth research questions address growing threats to organisations e.g., ransomware and DDoS attacks (Chapters 5 and 6). The options available to organisations once hit by such attacks and a stakeholder analysis of same is absent from the ethical literature. The analysis in this thesis highlights that both attack vectors are growing threats to organisations whilst examining how stakeholders' interests may be positively or negatively impacted by organisational responses to both threats.

1.4 Outline

Chapter 2 provides a detailed account of the results of a literature review that reveals the ethical issues in cybersecurity as identified in the academic literature. This chapter highlights the most discussed ethical issues in the ethical literature and identifies blind spots that the literature overlooks. Chapter 3 describes Robert Phillips' stakeholder theory that can be applied to complex organisational decision-making. Research questions two, three and four directly align with the blind spots revealed in the literature review. The analysis of these research questions can be found in Chapters 4, 5 and 6. Chapter 7 comprises a summary of the results of this research, discusses its limitations and offers suggestions for future research.

Chapter 2 Ethical Issues in Cybersecurity in the Business Domain

2.1 Introduction

This chapter is based on a published literature review included in the White Paper 'Cybersecurity and Ethics' (Yaghmaei, et al., 2017). The white paper is divided into three sections with each section covering ethical issues in cybersecurity in one of three domains: health, business, and national security[9]. The author co-wrote the business section of the white paper and a substantial amount of material from the author's research is included in this chapter. The aim of the white paper was to analyse and discuss current ethical literature on cybersecurity in health, business, and national security. One compelling point unveiled by the White Paper is that 'Ethics and Cybersecurity' is not an established subject, neither academically nor in any other domain of operation. In fact, it is a rather underdeveloped topic within Information and Communication Technology (ICT) ethics, where most published work focuses on privacy or the ethical issues associated with surveillance (Yaghmaei, et al., 2017:3).

The author's contribution to the business section of the literature review is discussed in this chapter because the work of the author (1) highlights the ethical issues that are described by the ethical literature in relation to cybersecurity in business and (2) illustrates that there are topics that have been overlooked by the ethical literature but are worthy of further ethical scrutiny. The healthcare and national security sections of the White Paper are not included in this chapter as the author did not directly contribute to those sections. In addition, it is the intention of the author to focus specifically on business organisations (as opposed to government and public sector organisations).

2.2 Methodology

The below methodology for the literature review was developed, agreed upon and executed by the author of this thesis with the help of the Coordinator of the CANVAS project and his post-doctoral fellow.

The author compiled and sent a list of key words which pertain to cybersecurity, ethics, and business to the post-doctoral fellow.[10] These words were used by the post-doctoral fellow to try to characterize

[9] The White Paper was a group effort. The methodology for each literature review was collectively agreed upon and worked on by a number of members of the project, namely the Coordinator of the CANVAS (Constructing an Alliance for Value-driven Cybersecurity) project, his post-doctoral fellow and the individuals involved in each reference domain. The author of this thesis was the lead researcher and author for the business domain.

[10] The following list of key words which pertain to cybersecurity, ethics, and business was sent to the post-doctoral fellow. The keywords for "Cyber" include: "Computer Security" OR "Cyber Security" OR "Cybersecurity" OR "Cyber-security" OR "Data Security" OR "Hardware Security" OR "Information Security" OR

the field-specific Boolean search expressions to allow for the identification of sources that discuss ethics, cybersecurity and business in any significant way. The keywords were used to conduct initial searches in the Web of Science (WoS) database and Scopus databases. Other exclusion criteria include sources written in a language other than English and sources published before 1996. The results yielded in excess of 6400 sources for the business domain. This number was too large for a search by hand. With the aim of searching approximately 1000 papers by hand, it was deemed necessary by the author and project team to refine the search items further.

The IT and ethics experts contributing to the white paper agreed on the basis of their expertise and previous exposure to cybersecurity issues in business that it would be very unlikely to find technical papers that discuss ethics, cybersecurity, and business in any significant way. Consequently, sources published in technical journals were excluded. Only non-technical subject categories were used. The following WoS subject categories were used: management, business, operations research, management science, ethics, social issues, humanities multidisciplinary, history philosophy of science, social sciences interdisciplinary, women’s studies, sociology, law. The following subject categories in Scopus were used: business, decision sciences, art and humanities, social sciences, economics, econometrics and finance, undefined[11].

After removing the technical sources, the author of this thesis conducted a manual review of the remaining titles. Any sources that appeared to not pertain to ethics, cybersecurity and business in any significant way were excluded. This process reduced the number of sources to 1451.

"Internet Security" OR "IT Security" OR "Mobile Security" OR "Network Security" OR "Security Breaches" OR "Security Of Data" OR "Security Requirement*" OR "Security Software" OR "Security System*" OR "Security Threat*" OR "Security Vulnerabilit*" OR "System Security" OR "Web Security" OR "data leak*" OR "non-repudiation" OR “sigint” OR "voting system" OR “cryptography” OR “cyberattack” OR "cyber attack" OR “cyberconflict” OR "cyber conflict" OR “cyberdefense” OR "cyber defense" OR “cyberterrorism” OR "cyber terrorism" OR "cyber threat*" OR “cyberthread*” OR “cyberwar*” OR "cyber war*" OR "computer crim*" OR "cyber crim*" OR “malware” OR “firewall” OR “botnet*” OR "denial of service" OR DDoS. The keywords for “Business” include: “banking“ OR “business“ OR “commerce“ OR “company“ OR “consumer“ OR “finance“ OR “payment“ OR “sales“ OR “shopping“. The keywords for “Ethics” include: “autonomy“ OR “privacy“ OR “value-driven“ OR "value driven" OR "European Value*" OR “value-profile“ OR “ethic*“ OR “responsibilit*“ OR “accountability“ OR “right*“ OR “value-sensitive“ OR "value sensitive" OR “moral*“ OR "informed consent" OR “philosoph*“ OR “equality“ OR “freedom“ OR “ethic*“ OR "contextual integrity" OR “politic*“ OR “dignity“ OR “democracy“ OR “discrimination“ OR “unfair*“ OR “fair*“ OR “non-discrimination“ OR “utilitarian“ OR "diversity issue*" OR “trustworthiness” OR “transparency“ OR “confidentiality“ OR “accountability“ OR “voluntariness“ OR “accessibility“ OR “justice“ OR “diversity“.

[11] The number of entries per category was checked after this refinement. If the number was <70, all entries were taken. If the numbers were >70, only the 50 most cited papers were taken. The reason for this strategy was to identify candidates potentially relevant for the ethics of cybersecurity. A cut-off value for a minimal number of citations was also applied per paper. The cut-off value was 6 citations for WoS and 9 citations for Scopus. Furthermore, as the citation criterion includes a bias for older papers (where more time was available to generate citations), the first 500 papers (in terms of publication date) were chosen.

The author of this thesis manually reviewed the 1451 sources by title and abstract. In the absence of an abstract or if no abstract was available, the first page of the source was reviewed. The sources were classified into the following three categories: (Category A) sources that *discuss ethics, cybersecurity and business in a significant way*; (Category B) sources that named and superficially mentioned either an ethical value or theory but upon review, *did not provide any substantive content regarding ethics, cybersecurity and business*; and (Category C) sources with a title suggestive of ethical content (e.g., The ethics and law on privacy enhancing technologies) but upon review, *did not provide any content relating to ethics, cybersecurity and business*.

Many sources had misguiding titles which are suggestive of ethics, cybersecurity and business (Category B or C). Upon review of the abstract and/or the first page of the source (when abstracts were unavailable), the majority of the sources offered little insight (Category B) or no insight (Category C) to ethics, cybersecurity and business. This means that Categories A and B are linked as they both mention cybersecurity, business and an ethical value or theory in the title or abstract, but they differ as regards substantiveness and relevance to ethics, cybersecurity and business.

Of the 1451 papers, a mere 23 sources were found to discuss ethics, cybersecurity and business in any significant way (Category A). A full text review of all the sources in Category A was conducted.

Snowballing by title was undertaken by checking the bibliographies of all 23 sources in Category A for any potentially relevant references. An abstract review of any titles that appeared to discuss ethics, cybersecurity and business in any significant way was undertaken. Where no abstract was available, a first-page review was completed. 6 additional sources were discovered in this way and were included in the final analysis.

The wider CANVAS Project team members agreed that any relevant papers that were not found from the Boolean searches but were considered relevant literature, discovered either through conference visits or through professional exchanges with colleagues, should be included. As a result, 4 additional sources were included.

This methodical approach yielded 33 eligible sources for the business domain. A full review of the 33 sources was completed by the author of this thesis and the relevant content was included in the business section of the white paper.

In preparation for this chapter, the 33 sources were consulted for a second time by the author of this thesis in 2018. The aim of this undertaking was to review the 33 sources with fresh eyes. Where possible, the intention was to also provide more context to the reader of this thesis as regards the different ethical issues that arise in cybersecurity within the business domain and how those issues are described by the ethical literature.

2.3 Results

The literature review highlights that the two main cybersecurity threats to businesses are people and technology (F-Secure, 2018). It suggests that the majority of cyberattacks in business fall into one of two categories: opportunistic (45%) or targeted (55%) (F-Secure, 2018). Opportunistic attacks are when the perpetrator attacks a target simply because an opportunity to do so presents itself. Targeted attacks are where a specific target is selected, and a concerted effort is made to compromise that target.

The literature describes several preventative methods that organisations are adopting to reduce the likelihood of being successfully attacked. These methods include encryption techniques, privacy enhancing technologies (PETs), ethical codes of conduct, and general deterrence[12] techniques. These preventative methods do not offer businesses complete protection from adversaries as the number and severity of cyberattacks continue to exponentially increase year on year (Calyptix Security, 2017). Organisations need to prepare for the inevitability of an attack occurring as it is not a matter of *if* an attack occurs, it is a matter of *when* it occurs (ENISA, 2018; Kaspersky Lab, 2018 ; PECB, 2017). To prepare for such an eventuality, organisations are encouraged to have emergency response plans in place to reduce collateral damage.

The ethical literature also indicates that there are a number of ethical issues that arise in cybersecurity (See *Table 1*). Ethical issues in cybersecurity do not always arise in the same context. For example, privacy is raised as an ethical issue in cases where organisations are secretly monitoring employees in the absence of consent. Privacy is also raised as an ethical issue in relation to organisations using a Cloud Service Providers (CSP). The former relates to invading the privacy of employees. The latter alludes to reducing organisational costs at the expense of increasing cybersecurity risk which can negatively impact the privacy of customers if the organisation were to fall victim to a cybersecurity breach.

[12] Deterrence is the prevention of action by either the existence of a credible threat of unacceptable counteraction and/or belief that the cost of action outweighs the perceived benefits (McKenzie, 2017).

2.3.1 Cybersecurity threats to businesses

This section includes 1) a list of the main cybersecurity threats to businesses as described by the ethical literature and 2) a discussion on how organisations are responding to those threats.

1) The main cybersecurity threats to organisations originate from a) malicious employees and b) cybercriminals exploiting weaknesses in people and technology (F-Secure, 2018)[13].

a) Malicious employees are considered one of the main cybersecurity threats to a business (Gunarto, 2003; Leiwo & Heikkuri 1998; Posey, et al., 2011). They can be employees who are motivated by job dissatisfaction, greed, pressing financial problems, political or social activism motives or may seek to compromise client data for financial gain (Leiwo & Heikkuri 1998; Simshaw & Wu, 2015;). Incident reports suggest that insider attacks account for between half and three quarters of all security incidents (D'Arcy & Hovav 2009; Da Veiga, 2016) and are the most significant threat to cybersecurity in respect of data leakage (Da Veiga, 2016).

b) Cybercriminals successfully target organisations by exploiting two main weaknesses, (i) people and (ii) technology.

(i) People: 52% of external attacks emanate from cybercriminals using sophisticated social engineering techniques (both targeted and opportunistic) to gain unauthorized access to businesses (F-Secure, 2018). They do so by exploiting weaknesses in people and manipulate victims into divulging information or performing actions that aid them in meeting their ends. Well-known examples of social engineering include tricking victims into installing malware (via email attachments or web links) and fooling users into sharing credentials (via fake login pages).

(ii) Technology: 48% of external attacks are from cybercriminals attacking a company's technical weaknesses (F-Secure, 2018). These attacks are both targeted and opportunistic and mostly involve an exploit against an unpatched vulnerability. Attacks targeting unpatched vulnerabilities are generally very prominent in the weeks following the disclosure of a vulnerability/exploit in a popular piece of server-side software. This is comparable to announcing to burglars that the master key for the local bank is kept in the plant pot beside the front door of the building. In other words, publicising unpatched vulnerabilities provides key information to attackers on how to gain unauthorised access to valuable systems or data.

[13] F-Secure is one of the CANVAS team partners. When they updated their latest Incident Report in 2018, the details of same were updated in this chapter.

In the proceeding sections, it is discussed how organisations are choosing to respond to a) malicious employee behaviour and b) cybercriminals exploiting weaknesses in people and technology in different ways.

a) Malicious employees:

The literature suggests that managing the insider threat not only involves implementing anti-malware technology, but also requires the appropriate management of people such as fostering a supportive environment (Kouatli, 2016). Traditional ways to block negative employee behaviour involve implementing technical measures such as authentication and identification, passwords and pass phrases, firewalls, intrusion detection, rights management, countermeasures, and system controls (Lowry, et al., 2014). Additional approaches include policies and procedures relating to employee misconduct, computer monitoring, audit trails, IT audits, information security (IS) risk analysis, IS countermeasures and general violation-prevention strategies (Lowry, et al., 2014).

In terms of policies and procedures, businesses use ethical codes of conduct to counter the insider problem despite codes having received criticism. For example, the codes are allegedly being used as a public relations gimmick or a means for protecting the corporation from legal liability, they lack impact and are nothing more than pseudo-ethics as they codify existing rules and standards of behaviour and do not encourage ethical reasoning when an individual is faced with new or difficult issues such as those which confront is personnel (Harrington 1996). Brey (2007) questions the effectiveness of ethical codes of conduct as, in general, they do not offer any details on the moral dimensions of security issues and how to cope with them.

The International Federation of Information Processors (IFIP), the Association for Computing Machinery (ACM), the Institute of Electrical and Electronic Engineers (IEEE), the British Computer Society and the Institute of Data Processing Management (IDPM) have all recognized the need for new codes of ethics to inform and advise their members about relevant social and ethical issues in cybersecurity (Gunarto, 2003). Others suggest that adopting codes and practices which include psychology methods such as using fear appraisals, general deference theory or related penalty-oriented techniques and/or leveraging employee perceptions of it policies can result in the policies appearing more mandatory (lowry, et al., 2014). Lewio & Keikkuri (1998) note that while general deterrence techniques decrease computer misuse, there is little evidence that they reduce the number of insider offences.

In respect of computer monitoring, this is viewed as a more intrusive approach to tackling the insider problem. As previously noted, some businesses conduct secret surveillance of employees with one study revealing that 21.6% of corporations search employee files (emails, network messages, voicemail) on the authority of executive managers, and in 66.2% of such cases, employees are not warned (Leiwo & Heikkuri, 1998).

b) Cybercriminals exploiting weaknesses in people and technology:
Improving employees' cybersecurity knowhow is one of the main ways that organisations are trying to manage the problem of cybercriminals gaining access via spam emails and social engineering (F-Secure, 2018).

To grapple with the issue of cybercriminals exploiting vulnerabilities in technology, organisations are conducting regular vulnerability testing on their systems. this can be done by upgrading firewalls to improve network resilience, improving anti-malware protection and the automation of patch management (Lowry, et al., 2014). pets can also be used as technical ways to protect personal identity specifically those that involve encryption in the form of digital signatures, blind signatures, or digital pseudonyms (Walters, 2001). Walters (2001) argues that these technologies may promote and protect privacy and security rights. Subsequently, Walters (2001) suggests that smart cards and biometric technologies can utilise pets in ways that protect privacy, human freedom, and well-being.

An alternative way to fortify one's defences against cybercriminals exploiting weaknesses in people and technology is to adopt new innovative technologies such as biometric technologies. Biometrics can identify or verify someone's identity based on physiological or behavioural characteristics (brey, 2007). For example, a person can be recognised by traits such as fingerprints, hand geometry, signature, retina or voice. It can be a reliable method of access control and personal identification for organisations such as financial institutions (Venkatraman & Delpachitra, 2008).

2.3.2 What are the ethical issues in cybersecurity in the business domain?

This section includes an in-depth discussion of the ethical issues as described by the ethical literature in relation to cybersecurity. The author conducted a second review of the 33 ethical sources used in the white paper "Ethics and Cybersecurity" which provided the author with a more succinct outlook on the frequently mentioned ethical issues in the literature. The ethical issues are tabulated and numbered from 1-15 in order that they arise in the literature. The most frequently mentioned issues are listed at the top of the table and the least frequently mentioned issues are towards the bottom (See Table 1). The issues range from privacy and accessibility to autonomy, ownership and usability. It is important to note that the literature consulted differentiates general ethical issues in cybersecurity from those in the cloud and from those in data sharing environments. This is because these two services expose information and data to new threats. For this reason, the next section brings context to the way in which the ethical issues listed in Table 1 arise.

Privacy is the most popular issue raised in the ethical literature, occurring in 27 of the 33 articles consulted (Abreu, et al., 2015 and 2016; Alouane & Bakkali, 2015; Bennasar, et al, 2015; Bodle, 2011; Brey, 2007; Conger, et al., 2013 Dean, et al., 2016; De Veiga, 2016; Dhillon, et al., 2016; Dodig-Crnkovic, 2004; Gunarto, 2003; Kouatli, 2016; Matwyshyn, 2010; McReynolds, 2015; Pearson, 2013; Posey, et al., 2011 Rifaut, et al., 2015; Robertson, et al., 2010; Salman, et al., 2013; Taddeo, 2013 and 2015; Venkatraman & Delpachitra, 2008; Walters, 2001). Privacy arises in relation to the secret surveillance of employees. One 2007 study showed that nearly half of all organisations surveyed monitor employees' computer activities without their consent (Posey, et al., 2011). Monitoring employees is a method adopted by organisations in an attempt to combat the insider threat. An insider threat is a cybersecurity risk to an organisation that stems from within the four walls of the organisation. It pertains to any individual or group of individuals who cause accidental or intentional harm to the organisation. However, secretly monitoring employees in the absence of consent to reduce the threat of the insider problem could be perceived as an infringement of employees' privacy and **autonomy** (Brey, 2007).

Table 1

	Ethical Issues in Cybersecurity in Business	Number of times identified in the literature
1	Privacy	27
2	Protection of data	26
3	Trust	23
4	Control	20
5	Accessibility	19
6	Confidentiality	18
7	Responsibility on businesses to use ethical codes of conduct	15
8	Data Integrity	14
9	Consent	12
10	Transparency	11
11	Availability	9
12	Accountability	9
13	Autonomy	8
14	Ownership	6
15	Usability	1

Table 1 showcases the ethical issues in the order of frequency in which they occur in the ethical literature.

Schoeman argues that a person has privacy to the extent that others have limited access to information about him or limited access to the intimacies of an individual's life, or limited access to a person's thoughts or their body (Schoeman, 1984). Brey (2007) makes specific reference to technology interfering with individuals' privacy. He argues that privacy entails securing the processing of personal information, including technologies that may observe and interfere with human behaviours and relations and their body, and their personal belongings (*ibid*.). He adds that a threat to privacy can include anything from defamation, harassment and manipulation to blackmail, theft, subordination and exclusion. In addition, Brey (2007) asserts that data breaches can cause psychological harm to a data owner if the information has been misused, lost, or stolen and is considered valuable, private or confidential to the data owner. While Walters (2001) argues that a threat to privacy is a threat to personal integrity.

Privacy is raised in the context of who can reduce costs and increase profitability for an organisation by improving access to data. CSP improve data **availability** through a process called replication. Replication is the process of copying original data and saving and storing it to several servers located across different jurisdictions. Such data can include customer or payroll information, suppliers' details, financial transactions, intellectual property, and trade secrets. Replicating data and storing it in different locations means that there is no single point of failure. If one server goes down or is compromised, the data can be accessed elsewhere. In this way, the demand for constant availability is met, yet it comes at the cost of being constantly vulnerable to attack. This is a move away from traditional business processes where organisations may have been vulnerable to an attack during business hours. Take a bank for example. Traditionally, a bank would open for a set period of time on specific days of the week. During opening hours, the risk of being attacked greatly increases because the front doors of the bank are open to the public, allowing any potential bank-robber to freely enter the building. In cyberspace, if data can be accessed around the clock, it is constantly vulnerable to an attack. Consequently, data that is processed and stored in the cloud is known for having a higher risk of being lost, stolen or misused (Abreu, et al., 2015 and 2016; Alouane & Bakkali, 2015; Bennasar, et al, 2015; Bonner & O'Higgins, 2010; Brey, 2007; De Veiga, 2016; Kouatli, 2016; Pearson, 2013).

Privacy is also discussed in respect of grey hat hacking. Grey hats are a variant of ethical hackers[14] whose sole purpose is to improve the security of cyberspace by finding security vulnerabilities in systems and networks. The process involves gaining unauthorised **access** to systems and data. Despite grey hats having good intentions, their endeavours are undertaken in the absence of systems owners' **consent** (Alouane & Bakkali, 2015; Brey, 2007; Dean, et al., 2016; Dodig-Crnkovic, 2004; McReynolds, 2015; Pearson, 2013; Posey, et al., 2011; Salman, et al., 2013; Shakib & Layton, 2014; Simshaw & Wu, 2015; Walters, 2001). Grey hats believe that they are improving the security of cyberspace by gaining unauthorized access to systems, finding vulnerabilities, and forcing systems owners to fix them. The fewer vulnerabilities there are in cyberspace, the safer it becomes for all. However, grey hats' work

[14] Ethical hacking involves breaking and entering systems, either authorised (usually a white hat) or unauthorised (commonly a grey hat) with the sole intention of fixing (or "patching") the point through which entry was gained.

involves invading the privacy of others and is undertaken in the absence of consent (Brey, 2007; Leiwo & Heikkuri, 1998; McReynolds, 2015). According to McReynolds, this practice is not well received by Western societies where high value is placed on privacy and intellectual property (McReynolds, 2015). Bypassing consent is an infringement of **autonomy**.

Another ethical problem that originates from grey hat hacking is **trust**. Stumbling across private data is not uncommon when ethical hackers search for vulnerabilities in information systems and networks. What happens in cases where security experts accidentally discover information that is suggestive of unethical or criminal behaviour? Grey hats and in particular white hats (white hats are ethical hackers who have been given the authority by the systems owner to test for security vulnerabilities) must decide where their obligations lie, i.e., do they have an obligation to respect the privacy and confidentiality of the systems owners and focus on finding security vulnerabilities only? Or should they notify the systems owner? Or directly contact the relevant authorities of the privileged information discovered? It could thus be advantageous to pre-empt such as situation prior to employing white hats and have in place a protocol regarding the accidental discovery of unfavourable information.

The **protection of data** from data breaches (where data is considered any information held within technology) is the second most frequently-mentioned ethical issue in the literature (Abreu, et al., 2015 and 2016; Alouane & Bakkali, 2015; Bennasar, et al, 2015; Brey, 2007 Conger, et al., 2013; Dean, et al., 2016; De Veiga, 2016; Dodig-Crnkovic, 2004; Gunarto, 2003; Harrington, 1996; Kouatli, 2016; Leiwo & Heikkuri, 1998; Matwyshyn, 2010; McReynolds, 2015; Pearson, 2013; Pieters, 2011; Posey, et al., 2011; Rifaut, et al., 2015; Robertson, et al., 2010; Shakib & Layton, 2014; Simshaw & Wu, 2015; Venkatraman & Delpachitra, 2008; Walters, 2001). Data breaches are escalating, not only in frequency, but also in severity (Matwyshyn, 2009) and can cause serious economic losses for a business (Brey, 2007). For example, Sony Corporation has been hacked over forty times since 2002, which has caused them an estimated global loss of over 100 million consumer records (Matwyshyn, 2009). The average cost of a data breach rose from 4.8 million USD in 2006 to 6.3 million USD in 2007 (Matwyshyn, 2009).

A failure to protect data can not only have a direct financial cost on organisations, but it can also cause irreparable damage to the victim organisation's reputation. The National Survey on Data Security Breach Notification (2005) finds that 60% of consumers are likely to discontinue their relationship with a firm after a data security breach, even if they did not directly suffer from the breach (Robertson, et al., 2010). From the end-user’s perspective, protecting personal information is of paramount importance. The user imparts a level of **trust** in a business when they share their personal information and assume their data will be protected from loss, theft or modification. If their data is of personal, cultural or social value and is compromised, this can cause psychological or emotional harm to the data owner (Brey, 2007).

Matwyshyn (2009) elaborates on the potential harm resulting from data breaches and contends that companies have a duty to avoid knowingly causing harm to others. He states that organisations have a duty to exercise capabilities for the greater social good (*ibid*.). Brey (2007) argues that organisations have an ethical obligation to **protect data and personal information** due to the potentially serious consequences a data breach can have on a business and its stakeholders. The consequences of inadequate protection depend on the business sector as a data breach for one company may be more detrimental than another. For example, a law firm holds highly valuable information including corporate records and personal information relating to clients, intellectual property, and trade secrets. Inadequate protection of such information can lead to a serious data breach that may threaten the very survival of the firm. This substantiates a duty on lawyers to use reasonable and adequate cybersecurity measures to prevent unauthorised access to client data (Simshaw & Wu, 2015). It is important to note that organisations who process European citizens' information are compelled by GDPR to share data breach information with a supervisory authority within 72 hours of discovering that a data breach has occurred (European Parliament, 2016). Stiff penalties are imposed for non-compliance with GDPR which include fining the victim organisation a maximum penalty of €20 million, or 4% of annual global turnover, whichever is greater (European Commision, 2016).

Trust is the third most frequently cited ethical issue in the academic literature and is described as crucial for the success of interpersonal and organisational relationships (Walters, et al., 2001). Many authors argue that there is a general distrust between consumers and organisations who process personal information or use advanced cybersecurity technologies that collect end-user information (Abreu, et al., 2015 and 2016; Alouane & Bakkali, 2015; Bennasar, et al, 2015; Brey, 2007; Conger, et al., 2013; Dean, et al., 2016; De Veiga, 2016; Dhillon, et al., 2016; Kouatli, 2016; Matwyshyn, 2010; McReynolds, 2015; Pearson, 2013; Pieters, 2011; Rifaut, et al., 2015; Robertson, et al., 2010; Shakib & Layton, 2014; Simshaw & Wu, 2015; Taddeo, 2013 and 2015; Walters, et al., 2001). It is believed that distrust stems from a lack of **transparency** in how businesses collect, process, store and share customer information, coupled with the uncertainty surrounding the type of security technologies organisations are using to protect or process users' data e.g., Privacy Enhancing Technologies (PETs), data mining techniques, etc (Abreu, et al., 2015 & 2016; Alouane & Bakkali, 2015; Bennasar, et al, 2015; Bodle, 2011; Conger, et al., 2013; Dean, et al., 2016; Pearson, 2013; Salman, et al., 2013; Shakib & Layton, 2014; Walters, et al., 2001). A lack of transparency regarding who has **access** to data (Abreu, et al., 2015 and 2016; Alouane & Bakkali, 2015; Bennasar, et al., 2015; Bodle, 2011; Brey, 2007; Conger, et al., 2013; D'Arcy & Hovav, 2009; Dean, et al., 2016; De Veiga, 2016; Dhillon, et al., 2016; Gunarto, 2003; Leiwo & Heikkurri, 1998; Lowry, et al., 2014; Matwyshyn, 2010; Pearson, 2013; Posey, et al.,2011; Simshaw & Wu, 2015; Walters, et al., 2001;) and who **controls** and **owns** the technology is a concern (Abreu, et al., 2015 and 2016; Alouane & Bakkali, 2015; Bennasar, et al, 2015; Bodle, 2011;; Brey, 2007; Conger, et al., 2013; D'Arcy & Hovav, 2009; Dean, et al., 2016; Gunarto, 2003; Matwyshyn, 2010; Pearson, 2013; Pieters, 2011; Posey, et al., 2011; Robertson, et al., 2010; Simshaw & Wu, 2015; Taddeo, 2013 and 2015; Venkatraman & Delpachitra, 2008; Walters, et al., 2001)

Many industries including aerospace, education, health and fitness, pharmaceutical, insurance, food, marketing and advertising, retailers, banking, e-commerce businesses, and even law enforcement are using advanced technologies to collect, analyse and buy or sell user information as a commodity in the absence of informed **consent** (Brey, 2007; Bodle, 2011; Dean, et al., 2016; Dodig-Crnkovic, 2004; Pearson, 2013; Salman, et al., 2013; Shakib & Layton, 2014). The supply chain of information collection usually begins at virtual hotspots such as social network platforms like Facebook, Twitter, and Instagram, search engines such as Google, and any website that tracks or obtains your browsing activity, IP address or personal information. Organisations can track, save, aggregate and circulate user information to any third party who is interested in users' online behaviour i.e., what and who users' search for, what their interests are, their gender, their socio-political or marital status, race, religion, what websites or social media posts they share with others, comment on or "Like" etc. This information is useful when it has been subjected to either data analysis or data mining. Data analysts use business intelligence and analytics models to try to make sense of raw data (a process called data profiling).

Profiling can involve accessing and collecting large amounts of raw data (in the absence of informed consent), organising it by tagging the data with keywords, descriptions or categories, and testing the quality and accuracy of the data. Data analysts group individuals who have certain characteristics that are associated with other traits and create a profile that might be used for other purposes e.g., police profiling to find criminals or terrorists. Data mining is a form of profiling that uses algorithms (including machine learning, statistical analysis and modelling techniques) to identify patterns, trends and subtle relationships hidden within the data. Mining can be coupled with data matching to link information from different data sets, creating a more comprehensive data image of an individual. Mining and matching uncover unique insights into individual online and offline behaviour that would have otherwise been unknown and are utilised by organisations to develop more effective marketing strategies, increase sales and decrease costs (Brey, 2007; Bodle, 2011; Dean, et al., 2016; Dodig-Crnkovic, 2004; Pearson, 2013; Salman, et al., 2013; Shakib & Layton, 2014). For example, a database that logs geolocations i.e., google maps can indicate where the user lives, works, buys their groceries, or attends a gym. This data can be mined and matched with an individual's social media data, which can suggest much more about the individual's behaviour i.e. how often the individual checks their social media accounts, how long they spend online (during working hours and non-working hours), what days and times of the month they are most likely to spend money, who they interact with most, such as family or friends or work colleagues, whether they read or share news from reputable sources, “Like” or comment on fake-news stories, what topics of conversation or pictures grab the user’s attention, how responsive they are to certain advertisements, how fast they scroll, or how quick they read.

While this personal information is categorised to create individual profiles and is sold to interested third parties such as data brokers or advertising, marketing, or insurance companies, the ethical literature suggests that it is improbable that individuals are aware that their online behaviour is constantly being tracked and traced and freely passed from one entity to another (Shakib & Layton, 2014). Shakib & Layton (2014) argue that there is no **transparency** between consumers and businesses and this level of interoperability is an infringement of **privacy** and **autonomy** when data-owners are not fully informed as regards what is happening to their data. The point is also raised that profiling can inflict harm on individuals as it involves stereotyping behaviour based on limited information and general assumptions which can lead to inaccuracies or even discrimination.

Confidentiality is an ethical issue for businesses who process personal information and who are at risk of a data breach occuring (Abreu, et al., 2015 and 2016; Alouane & Bakkali, 2015; Bennasar, et al, 2015; Bodle, 2011; Brey, 2007; Dean, et al., 2016; De Veiga, 2016; Kouatli, 2016; Leiwo & Heikkuri, 1998; Matwyshyn, 2010;Pearson, 2013; Robertson, et al., 2010; Salman, et al., 2013; Shakib & Layton, 2014; Simshaw & Wu, 2015; Venkatraman & Delpachitra, 2008;). Confidentiality is breached if unauthorised persons access private or valuable information. The attackers could use the confidential information against the victim to blackmail them or defame them. If the confidential data is of value, e.g. a driver's license or passport, it could be used for misappropriation (the illegal use of property or funds). If the confidential infromation included trade secrets, it could be sold or shared with malicious actors, the public, or competitors, which could damage the victim's reputation and result in a loss of market share.

If confidential data is tampered with this can compromise the **integrity** of the data and may result in a further financial loss for the data owner (Abreu, et al., 2015 and 2016; Alouane & Bakkali, 2015; Bennasar, et al, 2015; Bodle, 2011; Brey, 2007; Dean, et al., 2016; De Veiga, 2016; Dhillon, et al., 2016; Kouatli, 2016; Matwyshyn, 2010; Pearson, 2013; Simshaw & Wu, 2015; Venkatraman & Delpachitra, 2008; Walters, et al., 2001). An enquiry into the extent of the damage incurred may be necessary for insurance claims purposes. Additionally, when data is compromised, accountability becomes an issue as it is not always clear who exactly is **accountable** for the data breach. Was the organisation negligent in updating their systems which resulted in the data breach? Was it a social engineering attack that that could have been avoided? Or did the attack emanate from a third party like a CSP (Abreu, et al., 2015 and 2016; Conger, et al., 2013; Rifaut, et al., 2015, Venkatraman & Delpachitra, 2008; Dean, et al., 2016; Alouane & Bakkali, 2015; Kouatli, 2016; Pearson, 2013)? And if so, who is accountable?

Depending on the specific data breached and modified, it may also trigger legal investigations regarding **the ownership of data** (Alouane & Bakkali, 2015; Conger, et al., 2013; Dhillon, et al., 2016; Pearson, 2013; Shakib & Layton, 2014; Venkatraman & Delpachitra, 2008;). For example, modifying pharmaceutical trade secrets and sharing them with the public could be a tactic employed by adversaries to sabotage the victim organisation. Furthermore, data modification will affect the accuracy and quality of the data. Low quality data can compromise the security of an enterprise as security is directly linked to the accuracy of data (Dhillon et al. 2016).

A number of sources describe the **responsibility** that businesses have **to adopt ethical codes of conduct** as a way of reducing the likelihood of a successful attack (Abreu, et al., 2015 and 2016; Brey, 2007; D'Arcy & Hovav, 2009; Dean, et al., 2016; Dodig-Crnkovic, 2004; Gunarto, 2003; Harrington, 1996; Kouatli, 2016; Leiwo & Heikkuri, 1998; Matwyshyn, 2010; Pearson, 2013; Salman, et al., 2013; Shakib & Layton, 2014;). For corporations, adopting codes of conduct and general deterrence techniques are ways that businesses try to encourage ethical behaviour in the workplace in relation to computer use (D'Arcy & Hovav, 2009). Codes of ethics can keep employees abreast of laws and regulations and clearly outline unacceptable or illegal behaviour. According to Kouatli (2016), corporations can construct rules of conduct and codes of ethics to clarify responsibility. Harrington (1996) argues that codes of ethics are a means of deterring unethical behaviour, can be a basis for internal sanctions and can thus affect an employee's intentions. In their absence, it is easier for perpetrators to rationalize irresponsible behaviour (*ibid*.). Assuming that such codes have an impact on the decision-making process of an employee, they can contribute to (i) increasing awareness that an ethical problem in fact exists, and a potential computer abuse can occur and (ii) aiding the employee in making a judgment about right and wrong by clarifying right or wrong behaviour regarding the abuse (*ibid*.).

One final ethical issue raised in the literature relates to balancing **usability** and security. This is challenging as the more secure something becomes, the less convenient it is to use. Equally, the easier it is to access, the more vulnerable it is to an attack. This creates a dilemma between the usability of a service and keeping information secure. The challenge lies in the fact that consumers like to use convenient and easy-to-access services but also want their personal information secure (Dhillon, et al., 2016).

2.4 Discussion

This review highlights and discusses the main cybersecurity threats to businesses and discusses how organisations are choosing to respond to those threats. It also lists and describes the ethical issues that arise in cybersecurity for organisations as described by the ethical literature. The ethical literature highlights various ethical issues can arise in different contexts in cybersecurity which is something that organisations should be aware of and aim to resolve. This is simply due to the fact that cybersecurity can act in two opposing ways: it can function as a promoter of the interests of individuals such as the interest in easy to access products and services. Or it can work to undermine individuals' control over their own personal information, discrimination, and a breach of privacy and confidentiality. With this in mind, five conflicts of interest have been identified that can arise in organisations in relation to cybersecurity.

2.4.1 Conflicts of interests

Surveillance versus privacy, autonomy & justice

As previously mentioned, businesses are monitoring employees in the interest of countering the insider threat. Surveiling employees can range from recording all incoming and outgoing phone calls, emails, and voicemails to installing desk sensors to determine how long employees spend at and away from their desk. These measures conflict with employees' interests if they are perceived as an infringement of workplace privacy, autonomy and distributive justice. For example, if surveillance is carried out in the absence of consent, this impacts employee autonomy. From a legal perspective, if employers choose to secretly surveil their employees, Gunarto (2003) makes the point that employees have limited protection as the law appears to support employers' rights to read electronic mail and other electronic documents of their employees. However, Gunarto suggests that simply because an action is considered legal, it does not mean that it is ethical. He states, "in this matter, the definitions of 'right' and 'wrong' are not clear" (Gunarto, 2003: 1). Brey (2007) is more forthright on the ethics of surveillance. He maintains that surveillance can violate the notion of justice as it is in opposition to employees' expectations of just and fair treatment in the workplace in that employees expect their interpersonal space to be respected, they expect their rights to be respected, and they expect to be treated with dignity.

An extension of traditional workplace surveillance is a method called dataveillance. Dataveillance entails the large-scale computerised collection and processing of personal data to monitor people's actions and communications (Brey, 2007). This technique not only records and processes static information about individuals, but it also records and processes actions and communications which can be extended to customers (customer surveillance). Brey cautions against using these methods as they too raise ethical concerns over consent, privacy, and justice (2007).

Interestingly, research by Posey, et al. (2011) suggests that increases in organisational monitoring may lower commitment and may increase workplace deviance on the basis that monitoring efforts can be perceived by employees as an intrusion of privacy. If privacy invasion violates employees' expectations about fairness such as being treated with respect and dignity and having rights to interpersonal space, employees will perceive their employer's monitoring efforts as breaches of procedural and distributive justice (Posey, et al., 2011). Posey et al (2011) suggest that if employers' efforts are perceived as such, it may cause employees to engage in counterproductive behaviour such as "cyber-loafing".[15] This suggests that computer monitoring may in fact promote the activity that it is trying to curb.

Biometrics versus privacy & security

A paradox exists at the heart of biometrics[16] as on one hand the technology can be a threat to privacy for consumers as it is a technology of surveillance. On the other hand, biometric technologies can be utilized as security mechanisms that protect consumer privacy (Walters 2001). A trade-off also exists between securing biometric technologies and their usability. Biometric technologies can often require users to update their biometric data regularly if fault tolerant procedures are not in place. This could be considered inconvenient to consumers and put them off updating their biometric data which can affect the accuracy and security of the technology (Venkatraman & Delpachitra, 2008). One undesirable effect of biometric technologies is that they can leave traces of consumers information everywhere. Tracking and tracing consumers daily activities could eliminate anonymity and pseudonymity if there is widespread use of biometrics (Brey, 2007).

Implementing biometrics comes with security risks. For example, changes in lighting and photo angles in facial recognition can affect the reliability of data and prevent access to valid users. Masking a finger to avoid a match in fingerprint technology can affect the validity of matching accuracy. Hijacking of contour data in palm scanning/hand geometry could affect confidentiality and privacy. An inability to execute liveliness testing in iris/retina scanning opens the potential to print iris patterns on contact lenses and signature recognition can threaten data accuracy and reliability due to variable trait data (Venkatraman & Delpachitra, 2008). There is a risk with privacy and confidentiality if biometric information is in widespread use and is stolen or misused. Thus, moderating the security of biometrics is not just an operational challenge for organisations but also an ethical challenge as organisations must balance any identified conflicts of interest (Venkatraman & Delpachitra, 2008).

[15] Cyber-loafing is the unauthorised personal use of the internet in the workplace.

[16] Biometrics is the use of technology that collects information relating to physical, physiological or behavioural characteristics of a person that provide unique identifiers such as facial images or fingerprints (European Commission, 2021).

Cloud access versus security, privacy, confidentiality, control & usability

Conflict arises between organisations striving to provide consumers with constant access to services by using CSP versus consumers' interest in keeping their information private, confidential, and secure. Using CSP can reduce operational costs and increase profits for an organisation as it can relieve an organisation's need for an in-house IT department (Alouane & Bakkali, 2015; Kouatli, 2016; Pieters, 2011). However, increasing access by outsourcing to CSP increases the risk of data breaches, data misuse and data loss. If a breach involves valuable information being misused, lost or stolen, this compromises the confidentiality and integrity of the data. Traditional offshoring or outsourcing requires a business to protect data outside the four walls of the organisation. When financial institutions outsource to a CSP, security and privacy are the main areas of risk as "ensuring physical protection of data at a foreign site is more difficult than doing so at a local site" (Robertson, et al., 2010: 173). Conflicts of interest thus arise between the organisation's interest in offering constant access to services and consumers who value safety and security, service quality and data security (Robertson, et al., 2010).

Van den Hoven argues that access to information in modern society has become a moral right of citizens in the information age because information has become a primary social good: a major resource necessary for people to be successful in society (cited in Brey, 2007). Using the cloud also comes with the issue of control over data processing as customers' data is processed remotely in unknown machines (Alouane & Bakkali, 2015; Bennasar, et al. 2015). Pearson (2012) argues that there needs to be an appropriate level of access control within the cloud environment where it can often be unclear who controls the information and infrastructure, or who owns it (Pieters, 2011).

Responsibility and accountability issues arise when it is unclear who takes responsibility for the maintenance and backup of the information held in the cloud if it is stored in multiple locations (Kouatli, 2016). As previously mentioned, due to the practice of data replication, locating the data can prove very difficult as the system in use may automatically replicate data to different locations all across the world. This raises further security, ethical, and potentially legal issues if data is lost or stolen in a country where legislation on data protection and information security is not as stringent as the host's country (Kouatli, 2016). Pearson (2013) offers a solution to the conflict by arguing that security need not suffer in moving to the cloud. Organisations can instead outsource to security experts who can provide sufficient protection.

Responsibility can also be a challenge in cloud computing in respect of identifying which parties are responsible for which aspect of security (Pearson, 2013). Pearson (2013) notes that a number of threats coincide with cloud computing such as the nefarious use of cloud computing, insecure open Application Programming Interfaces (APIs), malicious insiders, shared technology issues, data loss or leakage and account/service hijacking. If there are entities within the provider chain that have inadequate security mechanisms in place, this can hinder security and increase the chance of being attacked (Pearson, 2013). The potential damage caused by inadequate security in the cloud provider chain can thus be greater than non-cloud environments due to the scale of operation and the presence of certain roles in cloud architectures and the fact that data may remain in the cloud for long periods of time resulting in a greater exposure time for an attack (Pearson, 2013).

Poor data quality in the cloud conflicts with some organisations' and data owners' interests. For example, when data is of poor quality this has two implications; the first is that the security of an enterprise becomes compromised as security is directly linked to the accuracy of data (Dhillon, et al., 2016). This is not in the interest of the organisation or the data owner. The second is that usability of a system comes into question if the system and data therein are not useful or if the data is out of context. This typically results in a loss of ownership and very serious security problems (Dhillon et al., 2016). The quality of data and information security in the cloud according to Robertson, et al. (2010) can be viewed as value-based issues that can vary in their moral intensity and can have a significant effect on businesses and their decision-making process (Robertson, et al., 2010).

Pearson (2013) adds that problems in the cloud can be viewed as trade-offs between security, privacy, compliance, costs and benefits wherein trust and transparency play a significant role. Pieters (2011) states that the cloud changes the containment-based approach to information security and forces organisations to implement data-level security instead. This has been referred to as de-perimeterisation which Pieters (2011) describes as the fading of the boundaries of organisations and their information infrastructure.

Hacker ethics versus security & ownership of information

Hacker ethics, as described by Knightmare, advocates the unauthorised access of systems with the aim of developing a more secure cyber environment for all. Knightmare encourages hackers to: "Never harm, alter or damage any computer, software, system, or a person in any way" and if damage is done, the hacker should do what is necessary to correct the damage and prevent it from occurring again (cited in Leiwo & Heikkuri, 1998, p. 215). Conflict can arise between hacker ethics and the general notion of privacy and intellectual property. A traditional ethical hacker view is that all information should be free, that access to computers should be unlimited and total (Brey, 2007). This is a strict ethical hacker ethic which not all ethical hackers subscribe to today. Tavani argues against the traditional ethic saying that the idea of all information being free is in direct conflict with the value of privacy, integrity, and accuracy of information. He suggests that information cannot be free as this means it could be modified at will and this runs counter to the very notion of intellectual property (cited in Brey, 2007). In addition, this belief implies that creators of information have no right to keep information to themselves nor can profit from it. Value conflict also arises as many ethical hackers have a common belief that by gaining unauthorised access into a system, they are providing a good outcome for the information security community by identifying vulnerabilities before they are exploited by adversaries (McReynolds, 2015).

Access in data sharing environments versus privacy, confidentiality, autonomy, control & security

Technologies enable society to benefit from online social engagement via increased interoperability between businesses and individuals. However, the price of engagement is paid by the individual who is encouraged to completely surrender their personal privacy on the internet. This creates a conflict between privacy and security (Shakib & Layton, 2014). Simshaw & Wu (2015) argue that businesses who benefit from processing, storing and analysing personal information have an ethical obligation to adequately secure personal data as this satisfies the interests of data owners. This includes all businesses that utilise and benefit from data mining techniques such as financial services, consumer products, manufacturing, the pharmaceutical industry, technology/services, retail, telecommunications, energy, and transportation (Dean, et al., 2016).

When data mining technologies are used alongside APIs, this can have “unforeseen ethical consequences” on individuals’ autonomy and privacy (Gattiker & Kelley, 1999, p. 223). Consider the following example. Facebook use data mining techniques and API tools that enable their users to navigate from site-to-site. This means that users can “log in via Facebook” on many sites via Facebook’s API. This allows Facebook users to easily create “new accounts” with various entities at the click of a button. All the while Facebook tracks, traces and disseminates any personal information that they have on the individual user (including name, profile picture, gender, networks, user identification (UID), list of friends) with the website the user has logged into. Facebook also have the authority to share this information with articulated networks and with third-party sites and services (Bodle, 2011). Bodle (2011) argues that users are unaware of the information that is being collected and used. He states that in circumstances where users have not given their informed consent for organisations to collect and use of their data, this is an infringement of privacy and autonomy.

APIs are not the only software used in data sharing environments that create a conflict of interests. Facebook’s Open stream and Instant Personalisation Pilot Program are other examples. The former allows outsiders to access a user’s entire Facebook real-time activity stream. The latter allows third party access to members’ data from which third parties can tailor content to the user’s tastes. According to Bodle (2011), both tools require enhanced security measures such as authentication as they both monitor online movement (Bodle, 2011). Yet, it is unknown if Facebook provides these measures.

As previously mentioned, one purpose of data mining can be to stereotype whole categories of individuals (Brey, 2007). Conger, et al., (2013) argue that conflict arises when the data owner has not consented to this type of analysis of their data which impacts their autonomy. Bodle (2011) acknowledges that soliciting an end-user's data flows increases data portability and the process of tailoring personalised content e.g., in targeted advertising campaigns is drawn from the data collected directly from the individual. However, he makes the argument that extensions of the techniques used are at the expense of user autonomy (Bodle, 2011). Brey (2007) & Dodig-crnkovic (2004) reiterate the consent issue and argue that individuals should be informed in respect of how organisations store, use or exchange personal and private information. They say that by obtaining consent, this lends true to the principle that a person should not be used as an instrument for advancing some goal – instead they should be fully informed and have freely consented to engage in an activity wherein their interests are respected.

The consent issue is tied to the value of trust, which can be viewed as a consequence of progress towards security and privacy objectives as trust revolves around the "assurance" and confidence that people, data, entities, information, or processes will function or behave in expected ways (Alouane & Bakkali, 2015). When trust is undermined, a power struggle emerges wherein one party has more power than the other (Kouatli, 2016; Pieters, 2011). Conger, et al., (2013) argue that this is yet another reason for businesses to implement adequate cybersecurity measures that balance the corporate use of data with security.

To reduce the security risks in data sharing environments, Conger, et al., (2013) discuss developing technologies that constrain data life by restricting the integration of data and erasing data when it reaches an age threshold. Technical solutions are suggested such as the adoption of firewalls (Gunarto, 2003) and the use of cryptography (Lewio & Heikkuri, 1998). Conger, et al. (2013) say that businesses could use PETs that involve encryption as they offer the promise of anonymity and the secure transport of data over a network. Walters (2001) agrees that while PETs are one technical solution to increasing security and protecting privacy, they have their limitations. He argues that PETS do not protect all parties equally, they weaken interpersonal bonds between communities as they are based on anonymity and they may negatively impact the formation of identity as identity develops in relation to others, not by isolating ourselves from others.

Salman, et al., (2013) note that this secondary use of data creates transparency and autonomy issues. Dean, Payne & Landry (2016) suggest that businesses who use data mining techniques should work from the Golden Rule's perspective – one should do unto others as he would have others do unto him. In other words, if you would not like your data to be used in a specific way, do not use the data of others in the same way. They argue that data miners ought to consider three moral requirements: 1) that the actor treat all acted upon equally and in like manner to action he would accept 2) that the person acted upon be regarded as inherently valuable and not just as a tool to attain the actor's own ends, and 3) that freedom of the person acted upon be respected. In doing so, the first looks at how information is collected, stored online and/or shared with others, the second considers how the collection or use of data benefits the data subject and the third commands the data miner to acknowledge and respect the autonomy of others (Dean, et al., 2016)

2.4.2 Blind spots

In addition to the identified conflicts of interest, three "blind-spots" have been identified in the ethical literature. The blind-spots have not received any systematic ethical analysis in the literature to date but are relevant and topical issues in the cybersecurity domain. They are as follows:

1) *Ethical Hacking*

   Despite one paper mentioning the potential use of ethical hackers or hacking specialists to identify vulnerabilities in systems (Leiwo & Heikkuri, 1998), this topic is not explored in any significant detail. One paper briefly questions the helpfulness of ethical hacking but again, does not go into any specific detail (Brey, 2007). One other paper tries to justify the circumstances under which it might be ethical for non-state actors to hack governments and corporations for political purposes (referred to as hacktivism) (McReynolds, 2015). Yet, the different types of ethical hackers i.e., PEN testers or red teams who can be employed externally by an organisation, or red teams, are not mentioned in any significant way. Their aims and functions nor the benefits and potential harm they may cause to an organisation and their stakeholders is overlooked. In fact, there are minimal comprehensive insights in respect of organisations employing PEN testers or red teams as a countermeasure to cybercriminals attacking organisations.

2) *Organisational responses to cybersecurity attacks*

   D'Arcy & Hovav, (2009) consider how a large percentage of organisations are choosing to not disclose cybersecurity breaches to the public. This can mean that a significant number of cybersecurity attacks go undetected. Based on that assumption, it is likely that figures released by industry surveys regarding computer crime underestimate the actual frequency of data breaches (D'Arcy & Hovav, 2009). Harrington (1996) asserts that businesses do not report illegal activity to law enforcement or impose severe sanctions on computer abusers as reporting is shunned, prosecution is complex, detection is uncertain, conviction is rare and rewards such as golden parachutes and well-paid consulting jobs are made available to convicted computer criminals. He concludes that computer abusers are rarely caught or punished – a fact well-known by potential computer abusers (Harrington, 1996).

In respect of the options available to businesses when they have been attacked and whether those responses are ethically appropriate is absent from the ethical literature. For example, when organisations fall victim to a ransomware attack or a DDoS attack, what is the ethically appropriate response? There is no mention of the harm that can be caused to customers, the cybersphere and wider society if an organisation decides to pay a ransom or negotiate a lower fee. Nor is there a discussion about adopting active measures e.g., attempting a botnet takedown in response to a DDoS attack.

3) Ubiquitous computing

Ubiquitous computing (also known as the IoTs) is described briefly by Brey (2007). His paper explains that ubiquitous computing is a process that entails embedding microprocessors into everyday working and living environments in an invisible and unobtrusive way (Brey, 2007). It incorporates wireless communication and intelligent user interfaces that use sensors and intelligent algorithms for profiling. This advancement can involve recording user behaviour patterns and adapting to different situations. For ubiquitous computing to function, it requires possibly hundreds of intelligent networked computers that are aware of an individual's presence, personality and needs enabling the technology to perform actions and or provide information based on the perceived needs (Brey, 2007). Brey (2007) concludes that securing this technology and data from criminals while also endeavouring to protect the privacy of the individual may prove extremely difficult as dozens of smart devices record activity and are connected to the developer's computers as well as third parties.

There is no mention that the IoTs devices enable DDoS attacks. There is no discussion surrounding the fact that more and more insecure IoTs devices are sold year on year which is contributing to the severity and frequency of DDoS attacks. In the absence of arrest and prosecution for cybercrime, there is no business ethics assessment of whether organisations should or should take matters in their own hands by attempting a botnet takedown to trace the source of the attackers.

2.5 Conclusion

There are a number of blind spots in the ethical literature that are worthy of further ethical scrutiny; ethical hacking, organisational responses to cybersecurity attacks and ubiquitous computing (called the IoTs). Chapters 4,5 and 6 narrow the focus on these blind spots respectively. For example, Chapter 4 is an analysis of the ethics of employing red teams (ethical hacking), Chapter 5 is an analysis of responding to ransomware attacks (organisational responses to cybersecurity attacks) and Chapter 6 is an analysis of attempting a botnet takedown in response to a DDoS attack (organisational responses to cybersecurity attacks and ubiquitous computing). Robert Phillips' conception of Stakeholder Theory (described in the next chapter, Chapter 3) is used to analyse these blind spots.

Chapter 3 Ethical Theory

3.1 Introduction

In this chapter Robert A. Phillips' description of ST and the moral and political theory of John Rawls is discussed. Phillips combines both into a single theory of organisational ethics which is used to ethically analyse three research questions in the succeeding three chapters. To reiterate, the three research questions are 1) Is it ethically appropriate for organisations to employ red teams to find security vulnerabilities? (Chapter 4) 2) What is the ethically appropriate organisational response to a ransomware attack? (Chapter 5) and (3) Is it ethically appropriate for organisations to attempt a botnet takedown? (Chapter 6).

According to Phillips, looking at old European cities it is easy to see a transfer of social power between different institutions throughout history. The oldest and largest buildings are religious in nature i.e. churches and cathedrals. They stood as the most powerful institutions for hundreds of years. When liberal notions of enlightenment emerged, governments replaced religion as the most powerful institution occupying the second oldest and largest buildings. Arguably, the most powerful institutions today with the newest and largest buildings are corporations. Just like religious instutitions and governments, large organisations are faced with making tough choices and trade-offs that can influence the lives of milliions of consumers across the globe. As philosophers attempted to analyse and justify the power wielded by the state and its agents, the same must now be done for organisations. And if ethics is to become part of business conduct, it must knit into organisational life (Phillips, 2003b). Phillips' theory of organisational ethics is used to analyse three research questions as it provides a general explanation of the creation and existence of moral obligations within organisations and among stakeholders. Phillips combines ST and the moral and political theory of John Rawls into a single theory of organisational ethics. Let us first consider ST.

3.2 Stakeholder theory & principle of fair play

ST asserts that the central goal of business should be to create value and trade for all stakeholders (Freeman, 2010). This requires management to do more than attempt to maximise shareholder wealth. Rather, management must account for stakeholders' interests and consider those interests during the decision making process (Freeman, 1984). Organisations should aim to make decisions that advance the well-being of legitimate stakeholders (Phillips, 2003b: 217). Legitimate stakeholders are determined by Phillips' work on the principle of fairness (Phillips, 1997: 52) and stakeholder legitimacy (Phillips, 2003a). Phillips interchangeably refers to stakeholders' well-being as stakeholders interests and notes that neither can be determined in the abstract (Phillips, 2003b: 31). In other words, the interests of stakeholders and their well-being depends on the decision to be made and how the organisation is arranged. In an attempt to remain as true to Phillips' work as possible, the interests of stakeholders are considered in respect of the decision to be made. For the purposes of clarity, a sample scenario for each decision to be made is provided. For example, the issue of employing red teams is specific to an organisation that has the financial flexibility to afford to pay for red team services. The legitimate stakeholders in this scenario are determined using Phillips' method and reasonable assumptions are made about the interests of all legitimate stakeholders i.e. the growth and sustainability of the organisation, career progression, competitive prices etc.

3.2.1 Analysing Rawls' Principle of Fair Play

Phillips (1997) turns to John Rawls' (1964) research on the principle of fairness as the philsophical basis of obligations of fairness to specific parties in commercial settings. Rawls' principle of fairness was originally suggested as a grounding for political obligations. Rawls describes the principle as follows:

> "Suppose there is a mutually beneficial and just scheme of cooperation, and that the advantages it yields can only be obtained if everyone, or nearly everyone, cooperates. Suppose further that cooperation requires a certain sacrifice from each person, or at least involves a certain restriction of his liberty. Suppose finally that the benefits produced by cooperation are, up to a certain point, free: that is, the scheme of cooperation is unstable in the sense that if any one person knows that all (or nearly all) of the others will continue to do their part, he will still be able to share a gain from the scheme even if he does not do his part. Under these conditions a person who has accepted the benefits of the scheme is bound by a duty of fair play to do his part and not to take advantage of the free benefit by not cooperating" (Rawls, 1964:9-10).

Phillips analyses Rawls' principle of fair play from an economic perspective. He isolates six conditions from Rawls' principle and argues that only four are applicable to commerical exchanges and disregards two (Phillips, 1997:54-55). The six conditions are 1) mutual benefit, 2) justice, 3) benefits only accrue under conditions of near unanimity of cooperation, 4) cooperation requires sacrifice or restriction of liberty on the parts of the participants, 5) the posibility of free riders exists, and 6) the voluntary acceptance of benefits of the cooperative scheme. Phillips accepts conditions 1), 4), 5) and 6) and skips 2) and 3).

He reasons that mutual benefit is a widely accepted concept in economics. Such "benefit" does not need to be direct personal benefit to the cooperator. For example, an employee who is a parent can be an active participant of the cooperative scheme only to obtain benefit for a friend or child. In such cases, the engagement between the organisation and the cooperator can still be considered a scheme of cooperation that is mutually beneficial.

Phillips dismisses the notion that justice is a necessary precondition for the existence of fairness based obligations in an economic context. While he accepts that an extorted promise could be considered no promise at all, "it does not follow that consent to an unjust insitutition is coerced" (Phillips, 1997: 55). To explain his point further, Phillips provides an analogy of a promise to an unjust person. He writes, voluntarily making a promise to an unjust person does not disqualify the obligation to fulfill the promise. Whether such an obligation dictates action can depend on whether the obligation can be overridden by other moral considerations such as the duty to fight injustice. Regardless of the justness of the insitution, a voluntary promise incurs an obligation to fulfill the promise. The idea that a "requisite feature of an cooperative scheme is that it be relatively just" is thus rejected by Phillips (1997:55).

Phillips discards the third condition as he "doubts the necessity of this precondition for obligations of fairness" (Phillips, 2003b; 122-123). He states that "if all of the other necessary conditions obtain, it would still be unfair to take the benefits of a cooperative scheme without contribution" - even if the benefits can be obtained with only one half or two thirds of the people cooperating (Phillips, 1997: 55). In addition, if a group is unknown or delimited, how can one determine how close they are to unanimity?

The fourth condition, that cooperation requires sacrifice and contribution, is accepted. Cooperation in an economic context can entail contribution and restrictions of liberty. For instance, an individual restricts their liberty by choosing to become an employee of an organisation as they must arrive at a specific time, stay for a particular amount of time, fulfill certain duties and abide by the rules of the organisation. While the organisation benefits from the product of the employee's labour, the employee benefits from career progression, receipt of a salary etc. Commercial transactions thus involve contribution and sacrifice in a cooperative scheme in which obligations of fairness arise.

Phillips accepts the fifth condition as well. The existence of free-riders in economic transactions is widely acknowledged in economic literature. Phillips accepts this condition because if there was no possibility of free-riders existing, it would not make sense to create a principle of fairness in the first place. It is only when there is a chance of free-riders existing that, "obligations of fairness take on an even greater significance" (Phillips, 1997:55).

According to Phillips, the sixth condition, the voluntary acceptance of benefits of cooperative scheme, is vital to the existence of obligations of fair play. He declares that it is the "voluntary acceptance of the benefits of a scheme that actually creates the obligations" described (Phillips, 1997:56). While this condition is not explicitly mentioned in Rawls' principle of fairness, Phillips claims that it is present in theory and is vital to the viability of the principle (Phillips, 1997:56).

3.2.2 Phillips' Principle of Stakeholder Fairness

From his analysis of Rawls' work, Phillips concludes that commercial actions qualify as cooperative schemes, in which obligations of fairness to specific parties arise under certain conditions (Phillips, 1997: 57). These conditions can be found in Phillips' amended version of Rawls' principle of fairness below. Phillips initially called his amended version his "working definition of fairness" (Phillips, 1997:57). In his later work, he refers to it as the Principle of Stakeholder Fairness (PoSF), which he explains "provides an explanation of the obligations due stakeholders" (Phillips, 2003a: 26;). The PoSF is as follows:

> "Whenever persons or groups of persons voluntarily accept the benefits of a mutually beneficial scheme of co-operation requiring sacrifice or contribution on the parts of the participants and there exists the possibility of free-riding, obligations of fairness are created among the participants in the co-operative scheme in proportion to the benefits accepted" (Phillips, 1997:57).

3.2.3 The cooperative scheme, stakeholder status & stakeholder obligations

The PoSF is operative when there is 1) mutual benefit, 2) the cooperative scheme requires sacrifice or restriction of liberty on the parts of the participants, 3) there is a possibility of free riders, and 4) there is a voluntary acceptance of benefits of the cooperative scheme. Under these conditions obligations of fairness are created in proportion to the benefits accepted. Phillips (2003b) links his PoSF, stakeholder identification and obligations in the following excerpt:

> "At a minimum, stakeholders are those groups from whom the organisation has voluntarily accepted benefits. By doing so, the organisation has incurred obligations of fairness to attend to the well-being of these stakeholders – at least insofar as their well-being is affected by interactions with the focal organisation. This will typically include groups such as financiers, employees, customers, suppliers, and local communities" (Phillips, 2003b; 217).

Phillips states that "stakeholder obligations, and therefore stakeholder status are created when the organisation voluntarily accepts the contributions of some group or individual" (Phillips, 2003a:26). Such obligations are distinct from duties and basic human rights which simply exist by virtue of someone being human. He explains, "obligations of stakeholder fairness are additional moral obligations that are created based on the actions of the parties" which are created between individuals and organisations within the sphere of 'private associations'" (as opposed to at the level of the basic structure of society) (Phillips, 2003b: 26-27).

In commercial transactions, proportionality is vital to the principle of fairness such that obligations of fairness are proportional to benefits received (Phillips, 2003b:126). He advises that it is "particularly important in the context of establishing obligations among organisations and their stakeholders" as it prevents "one from being disproportionately obligated to a cooperative scheme while accepting relatively little of the benefit thereof" (Phillips, 2003b:126).

3.3 Normative, derivative and non-stakeholders

In any given situation, Phillips explains that there are three types of stakeholders. One group who management are morally obligated to consider in their decision making is called normative stakeholders. Another group, who management have no direct moral obligation to but should consider due to their ability to help or harm the organisation or its normative stakeholders, is referred to as derivative stakeholders. And a third group management do not need to consider in their decision-making process due to the unlikelihood of this group having any influence or impact on the organisation and its normative stakeholders. This group is called non-stakeholders. In the subsequent paragraphs, we take a closer look at this stakeholder allocation by Phillips.

3.3.1 Normative stakeholders

Phillips argues that normative stakeholders are "stakeholders to whom the organisation has a moral obligation, an obligation of fairness, over and above that due other social actors simply by virtue of their being human" (Phillips, 2003a: 30). There is no specific set list of normative stakeholders as this allocation depends on the organisation's mutually beneficial cooperative scheme and the decision to be made. Phillips suggests that we can assume that employees are a good example of a stakeholder who voluntarily accepts the benefits of a mutually beneficial scheme of cooperation that involves sacrifice and contribution. Employees are typically a participant of the scheme of co-operation and as such are owed obligations of fairness. Similarly, Phillips argues that the local community is a stakeholder that is typically a participant of the scheme of co-operation i.e. one that is mutually beneficial and involves sacrifice and contribution on the parts of the participants. This is on the basis that the local community can offer tax incentives to organisations to encourage them to set up within their locality. The local community might sacrifice property to the organisation and in return local businesses and services can benefit from increased activity to the area, job openings for locals etc. Local businesses can act as suppliers to organisations. Phillips maintains that suppliers are also a typical participant of the scheme of co-operation (Phillips, 2003b; 217). Likewise, shareholders are typically a participant of the cooperative scheme. Shareholders initially sacrifice capital from which the organisation benefits as they can use such funding to execute their function. Through the long term prosperity of the organisation, shareholders can gain a return on their investment. Lastly, customers are a typical participant of the cooperative scheme as prescribed by Phillips. The organisation benefits from customers investing their time and money in the organisation by purchasing their products or services while the customer benefits from the use of those products or services. Collectively, Phillips refers to these typical participants of the cooperative scheme as normative stakeholders. In this thesis I conduct a stakeholder analysis that lists normative stakeholders as those who typically fall within an organisation's mutually beneficial cooperative scheme i.e. shareholders, employees, customers, suppliers, and local communities *(Ibid).* The ethical analysis considers normative stakeholders' interests at length and how those interests might be affected when 1) an organisation decides to employ a red team to find vulnerabilities in people and technology (Chapter 4), 2) an organisation chooses to pay, not pay or negotiate a ransom post-ransomware attack (Chapter 5) and when 3) an organisation decides to attempt a botnet takedown (Chapter 6).

3.3.2 Derivative stakeholders

Phillips accepts that there are groups of stakeholders who do not fall within the cooperative scheme but merit management's consideration based on their ability to influence, help or harm the organisation and its normative stakeholders e.g., competitors, the media, activists or terrorists. Phillips explains that management have a derived moral obligation to attend to these powerful stakeholders in order to advance the interests of the organisation and its normative stakeholders. Phillips thus, refers to such groups as "derivative stakeholders" (Phillips, 2003a: 31). He provides an explanation through the following example. If a situation arises where an activist group or competitor are threatening the survival of the organisation, "managers should exert as much time and effort as necessary dealing with this threat" (Phillips, 2003b: 221). Spending a limited amount of time and resources attending to the demands or threats of derivative stakeholders is a justifiable trade-off that is likely to be accepted by normative stakeholders when it becomes clear to them that it is in their interests for management to adopt this approach (Phillips, 2003b: 222). In keeping with Phillips' approach, derivative stakeholders are identified dependent on the decision to be made and their interests are briefly acknowledged in the stakeholder analyses conducted in Chapter 4, 5 and 6.

3.3.3 Non-stakeholders

There are also non-stakeholder groups - those whose potential impact upon the organisation or its normative stakeholders is minimal or completely absent. The assignment of non-stakeholder status simply means that there is no moral obligation to attend to the well-being of these stakeholders. There is no derived obligation to consider the interests of these groups as their potential impact on the organisation or its normative stakeholders is negligible. It would be wrong to assume that the organisation has no moral relationship whatsoever with non-stakeholders. The organisation has a moral obligation to non-stakeholders by virtue of them being human to the extent that the organisation cannot violate basic human rights. However, the organisation does not need to attend to the well-being of these stakeholders.

3.4 Prioritizing stakeholders

Phillips states that "normative stakeholders take moral precedence over derivative stakeholders" (Phillip, 2003b; 220). This is based on the fact that "concern for derivative stakeholders is only justified by reference to normative stakeholders" (Phillip, 2003b; 220-221). Should a conflict of interest arise amongst normative stakeholders, Phillips encourages a form of prioritisation that is based on equity. He asserts that "voice and share – and therefore a sort of priority – should be based on contribution to the organisation" (Phillips, 2003b: 223). In other words, the more a stakeholder contributes to the organisation, the greater their voice and share should be (Phillips, 2003b: 223). This is consistent with the PoSF which states that "obligations of fairness are created […] in proportion to the benefits accepted" (Phillips, 1997:57).

However, this is a prescription that is difficult to apply. For example, how can one measure contributions of shareholder capital, employees effort or customer loyalty? Therefore, determining who amongst the categories of normative stakeholders contributes the most and thus, has the largest voice can be challenging. It is also a hard task to prioritise stakeholders interests when there is a conflict between obligations to two different normative stakeholders. For example, by fulfilling your obligation to one stakeholder, you might violate your obligations to the other, and vice versa.

Phillips recommends the work of Habermas on discourse ethics as a possible source of adjudicating stakeholder conflict (Phillips, 2003b: 153). According to Habermas, if conflict exists between extant norms, management can attempt to resolve the conflict by engaging in discourse with stakeholders. Habermas calls this communicative action which he argues is different from strategic action. Phillips concurs with the two types of action being different and explains strategic action is where one party seeks to influence the behaviour of another party through the threat of sanctions in order to satisfy the interests of one actor (Phillips, 2003b: 58). Putting moral restrictions on the open dialogue between conflicting stakeholders is necessary according to Habermas and Phillips. For example, Habermas states that the conversation should be free from "external or internal coercion" which he calls the "ideal speech situation" (Habermas, 1990: 58). Additionally, Habermas argues that an action can be only considered moral if it "meets (or could meet) with the approval of all affected in their capacity as participants in a practical discourse" (Habermas, 1990:66). Phillips states that "the discourse by which stakeholder norms are tested must have moral (i.e. communicative) restrictions rather than being merely strategic in nature" (Phillips, 2003b: 153).

While Phillips claims that communicative action with moral restrictions is the ideal way for management to adjudicate stakeholder conflict, it is not the only way to adjudicate conflict. He argues that it is possible for management to conceptually undertake a top-down approach and arrive at a similar outcome produced from open dialogue if the top down approach includes moral restrictions (Phillips, 2003b; 157). The moral restrictions he includes require management to 1) take action that is likely to reach the communicative assent of all stakeholders and 2) take action that supports the continuation of the cooperative scheme. In this conceptual approach, the affected stakeholders are not present to vouch for their own interests. Management consider their interests in their absence and make an assessment based on their moral imagination how stakeholders interests may be affected by the taking different courses of action. As Phillips puts it, a good stakeholder manager should seek through "moral imagination to place herself in a position where she can anticipate the demands, actions and needs of the myriad of stakeholder groups" (Phillips, 2003a; 37). Choosing an action that supports the continuation of the cooperative scheme seems like an obvious criterion for management to consider when trying to manage stakeholder conflict. As this top-down approach can occur in the comfort of manager's offices, I argue that it is like an abbreviated approach, taking less time and using considerably fewer resources when compared to organising meetings with stakeholders.

In this thesis, a top-down approach is used. Using this approach is the first necessary step to addressing the ethics of 1) employing red teams 2) responding to ransomware attacks, 3) attempting botnet takedowns. This conceptual approach contributes to the minimal ethical literature that covers ethics, cybersecurity and business. The results from this research may support decision-making for individual companies. It also has the potential to act as a springboard for further conceptual or empirical analyses e.g., communicative action by academics or professionals that involves direct stakeholder engagement. Should any conflict of interest arise amongst stakeholders, Phillips' top-down approach to determine which action supports the continuation of the cooperative scheme and is likely to achieve the communicative assent of all stakeholders will be applied.

Chapter 4 Employing a Red Team

4.1 Introduction

People and technology are the two main cybersecurity vulnerabilities in organisations. The aim of this chapter is to determine under what circumstances it is ethically appropriate for organisations to employ a red team to find security vulnerabilities in people and technology. As previously mentioned in Chapter 1, red teaming is an "authorised, adversary-based assessment for defensive purposes" (Sandia National Laboratory, 2011). The author attempts to make this determination by analysing red teaming and how it affects stakeholders interests by using Phillips' method described in Chapter 3. This chapter begins with a description of red teams, ethical hackers and adversaries and explains how their intentions differ from one another (section 4.2.1). Red teaming is explained in terms of where it falls within the different layers of organisational cybersecurity and how organisations can make the most out of a red team engagement (section 4.2.2). Insights regarding the aims and objectives of red teams and PEN testers are shared (section 4.2.3) along with an explanation of how red teams work and can help an organisation fortify defences against adversaries (section 4.2.4). The ethics of red teaming is then analysed (section 4.3) and a conclusion is made as to whether it is ethically appropriate for an organisation to employ a red team for the purposes of improving the security posture of the organisation (section 4.4).

4.2 Red teaming

4.2.1 Ethical hacking

Red teams are typically comprised of ethical hackers called white hats. Hacking is understood as a process that involves a person gaining entry to the computer of another person, and a 'hacker' is the person committing the act (Himma & Tavani, 2008). A hacker is traditionally a highly skilled individual who is competent in the areas of cybersecurity, computer systems and networks. However various archetypes of hackers have emerged since the term was coined in the 1960s by programmers at Massachusetts Institute of Technology (MIT) who manipulated code to do exactly what they wanted it to do (Granger, 1994). White hats are employed by organisations to defend against people, networks and applications. They are different to their ethical hacking cousin, the grey hat, as grey hats attempt to improve security by scouring systems for vulnerabilities without the system owner's consent and notifying the owner if any vulnerabilities are discovered.

Both white hats and grey hats work with or alongside organisations to try and improve the security of cyberspace. This sets them apart from other archetypes of hackers known as hacktivists and adversaries. For example, hacktivists are politically motivated IT experts who use the Internet to execute specific political actions of defence or resistance to organisational or governmental actions. One famous hacktivist group is called Anonymous, who take action that they believe is in the public's interest. To date, Anonymous have executed denial of service attacks and have obtained and leaked confidential information to the public (Jordan, 2017). Adversaries include technically adept malicious code writers who develop and spread malware or sell it as crimeware. Adversaries who buy crimeware-as-a-service are commonly referred to as script kiddies. Script kiddies launch their own cybersecurity attacks without requiring any prerequisite programming knowledge. The intentions of both malware authors and script kiddies range from a desire to take organisations offline, to extortion or notoriety. Red teams' intentions are polar to those of adversaries, as red teams work with organisations and their internal security team to fortify defences by identifying vulnerabilities and fixing them, rather than exploiting existing vulnerabilities to cause harm.

4.2.2 Advanced security

In order for an organisation to get the most benefit out of employing a red team, the organisation must advance through the lower levels of security first. According to Lee (2015) there are five hierarchical levels of cybersecurity. Architecture, passive[17] and active[18] defence form the foundational levels of cybersecurity and intelligence and offence constitute as more advanced levels of cybersecurity. Red teaming falls under the most advanced level of security, offence. To get the most out of a red team engagement, organisations ought to establish a baseline of security first by adopting passive and active defences first, and later consider using intelligence and offensive strategies like employing a red team. To establish the foundational levels of security, organisations can utilise the function of an internal security team called a blue team. Blue teams typically manage passive and active defences within organisations. They have an inside-out view of the organisation with access to business objectives and security strategies (Accenture Security, 2020). They adopt defensive security measures to maintain internal network defences against all cyber-attacks and threats e.g., introducing stronger password policies, ensuring conformity to security procedures, running network vulnerability scans and monitoring and reporting any suspicious employee behaviour (Atkinson, 2018). Blue teams are also involved in gathering intelligence on potential adversaries, their actions, capabilities, tactics and techniques.

Blue teams help organisations manage security vulnerabilities in technology by adding passive defence systems to the base level of security, architecture. This includes installing firewalls[19], anti-malware

[17] Passive defence measures are actions that do not involve a direct, automatic reaction during an attack (Crane, et al., 2013). For exmaple, running regular scans and monitoring network activity would be considered passive defence measures.

[18] Active defence measures are actions that involve direct reaction to an attack. Attempting a botnet takedown in response to a DDoS attack would fall into this category.

[19] Firewalls act as a shield preventing unauthorised users from accessing private networks and preventing malicious software from accessing a computer or network via the internet.

systems[20] and intrusion prevention systems[21] (Giacomello & Pescaroli, 2019). These systems provide consistent protection against or insight into threats that exploit vulnerabilities in people and technology without constant human interaction (Evans & Horsthemke, 2019). The blue team can fortify these defences internally by monitoring, responding and learning from adversaries internal to the network (Lee, 2015:10). They can also try to find solutions to identified problems, and restrict or shut down any indicators of compromise in people and technology,[22] which includes detecting and responding to simulated attacks from ethical hackers like PEN testers or red teams.

The blue team are also responsible for ensuring that software bugs are patched and up-to-date. Atkinson (2018) states that the source of many technical security vulnerabilities in organisations is poorly designed hardware and software products that vendors do not fully test or that fail to follow security by design principles.[23] In essence, this means that software is often deployed by vendors despite it containing security faults. Over time, vendors tend to discover new vulnerabilities in their products and deploy patches to users in order to fix the vulnerabilities in the form of a 'software update'. Failing to run software updates means that the vulnerabilities remain unpatched creating a larger window of opportunity (known as the 'attack surface') (Shearwater, 2017) for adversaries to attack the user (Oakley, 2018).

It is possible for software bugs to go undetected by software or security vendors and be exploited by an adversary before a patch is released to the public. Unknown vulnerabilities like these are called 'zero days' and their existence increases the medium through which organisations can be targeted (Atkinson, 2018). Zero days are highly likely to be successful due to defences not being in place thus making them a severe security threat to organisations. Therefore, it is crucial that software updates are installed as soon as they become available. However, running software updates can be a time consuming process and can require a high level of organisation which can be enough to deter and dissuade an organisation from updating their systems regularly.

[20] Anti-malware is a software tool designed to identify and prevent malicious software or malware from infecting a computer or system.

[21] Intrusion prevention systems (IPS) monitor and collect information about a network for potentially malicious content.

[22] There is also a yellow, purple, green and orange team. Yellow teams are the developers i.e. software builders, application developers, software engineers and system architects. They build and design software, systems and integrations to make organisations more efficient. Their main focus areas are functionality, applications and automation. Orange teams facilitate interaction and education, and purple teams' try to maximise the results of red team engagements and improve the blue team's capabilities. The Orange team works closely with the yellow team with the intention of making the yellow team more security conscious, providing education to benefit software code and design implementation. However, in practice, yellow, purple, green and orange teams are not as common as blue teams.

[23] Security by design principles are intended to ensure that networks and technologies are designed and built securely (National Cyber Security Centre, 2019).

The problem with not running software updates is that malicious attackers readily take advantage of vulnerabilities in outdated software by writing malicious code that targets and exploits them. Once the malware is written, the adversary can choose to sell the malicious code on the Dark Web[24], they can launch a targeted attack on a particular company or they might decide to launch a general attack themselves. A general attack that targets unpatched software vulnerabilities can affect many organisations and cause significant collateral damage. The WannaCry attack in 2017 is a prime example of a general attack that took advantage of a Windows vulnerability in outdated software. Despite Windows making a patch available by way of a software update, numerous organisations including the NHS, Renault, Deutsche Bahn, FedEx, Nissan and Hitachi (Inagaki, 2017; Thomas, 2017; Vigliarolo, 2017;) failed to update their systems and were successfully targeted by WannaCry ransomware (Inagaki, 2017). Employing a red team or PEN testers to test internal security defences for weaknesses or flaws that may have gone undetected or unnoticed by the blue team i.e. outdated software, backdoors or network vulnerabilities, can help organisations fortify their defences against potential cybersecurity attacks.

4.2.3 Red teams versus PEN testers

Red teams and PEN testers are often conflated as being one of the same, when in fact they can have different approaches, processes and objectives (Thompson, 2019). For example, red teaming includes penetration testing techniques as well as intrusion testing on physical and real-life cyberattacks (Brangetto, et al., 2015). Red teaming is a full attack simulation that focuses on all areas of a business (Nicholson, 2019). The engagements are designed to achieve specific goals such as gaining access to a sensitive server or business critical operation. Red teams are focused on emulating an advanced threat actor using stealth, subverting established defensive controls and identifying gaps in the organisations defensive strategy which can include breaching networks and systems, using social engineering tactics and gaining physical access to premises and devices (Nicholson, 2019).

PEN testing, on the other hand, is an assessment of whether certain networks, assets, platforms, hardware or applications are vulnerable to an attacker. The blue team is usually aware of the scope of the PEN testing being conducted. Therefore PEN testing is not typically focused on testing the ability of the blue team to detect and respond to a simulated attack. PEN testing very often involves manual or automated testing that can typically take anywhere between one and three weeks (High Bit Security, 2020a), the results of which provide the sponsor organisation with a snapshot of the security posture of the organisation. The cost of employing PEN testers can range between 2,000-100,000 USD (Hacken, n.a.; High Bit Security, 2020b).

[24] The Dark Web is a collective of hidden internet websites that are only accessible by a specific web browser like Tor ('The Onion Routing' project) (Kaspersky, 2021).

In contrast, red teaming is usually carried out over a longer period of time (between 10-12 weeks) with costs dependent on the duration, networks, IP addresses, applications, facilities involved etc (European Central Bank, 2018a). Red team testing is often undertaken without the blue team knowing in advance and the goal of the red team is to avoid being detected by the blue team. This not only tests the blue team's ability to detect and respond to a real cybersecurity attack, but also provides organisations with a better understanding of their ability to detect and respond to real-world attacks. Red teaming is considered the 'gold standard' way to find vulnerabilities to fix in organisations, however its ethical aspects remain an understudied area[25] of research despite it involving the manipulation of people and technology (Yaghmaei, et al., 2017). Accordingly, in the proceeding sections of this chapter, I solely focus on red teaming, explain how it works and systematically analyse the ethics of employing a red team.

4.2.4 How red teaming works

Reconnaissance

Based on the direction of the sponsor organisation, Red teams are authorised to conduct various types of hacking such as network, email, password or computer hacking (Saha, 2020). Gathering information about the target can entail on- and offline research including foot printing and reconnaissance. This process involves collecting as much information about the target system as possible, such as internet protocol (IP) address, domain name system (DNS) information, who owns the domain and how to contact them, the operating system (OS) used, employee email I.D, phone numbers (GreyCampus, 2020). This information can reveal system vulnerabilities that can be exploited. For example, Google hacking[26] is used to identify security vulnerabilities in web servers/applications, gather information about arbitrary or individual targets, discover error messages disclosing sensitive information, or discover files containing credentials and other sensitive data (Acunetix, 2020). Scanning techniques such as OS fingerprinting can be used to find more information about the type of hardware or software used on targeted devices. This information can reveal security vulnerabilities that can be exploited in an attack.

[25] Few commentators mention the importance of PEN testers undertaking ethical training (Fulton, et al., 2013; Brenner, 2014; Stahl, et al., 2014; Nicho & Khan, 2014) but none examine in any significant way how red teaming can affect stakeholders' interests or try to determine whether employing a red team is an ethically appropriate way for organisations to identify security vulnerabilities for the purposes of improving the security posture of an organisation.

[26] Google hacking is where the Google search engine is used to find sensitive information or vulnerabilities that may be exploited (Lubis, et al., 2011).

Social engineering

Red team testing is run in a live environment. Stiawan et al. (2017) argue that this is necessary in order to establish how vulnerable an organisation is to a malicious attack. To test vulnerabilities in people, red teams employ social engineering tactics as they can play a critical role in the discovery and exploitation of human-centric security weaknesses in organisations (Hatfield, 2019). Social engineering is "any act of using influence, manipulation, or lying, so that a person performs an action that may or may not be in its interest" (Moinescu, et al., 2019: 1). Social engineering attacks typically exploit the natural human tendency to trust. They follow a common pattern that can be segmented into four phases: 1) collecting information about the target; 2) developing a relationship with the target; 3) exploiting the information available to them and launching the attack; and 4) exiting without a trace (Salahdine & Kaabouch, 2019).

Social engineering attacks can be human, or computer based. The former involves direct contact between the victim and the attacker, e.g., physical contact, eye contact or voice interactions. This can include gaining physical access to confidential documents, shoulder surfing, dumpster diving, or pretexting (Wilhelm, 2013). Shoulder surfing is simply watching the victim enter a password or sensitive information. Dumpster diving can entail collecting information from an organisation's general waste collection bins. This can include obtaining discarded confidential documentation and old computer materials or drives. Pretexting is where the attacker fabricates a scenario in order to convince victims to share personal or confidential information (Wilhelm, 2013). Usually this involves the pretexter taking advantage of weaknesses in identification techniques in voice transactions. These attacks can take some time whereas computer-based attacks can be executed by red teams remotely through computers or mobile phones in a matter of seconds. Computer based attacks include SMS or email scam attacks, malicious pop-up windows or ransomware (Salahdine & Kaabouch, 2019).

Maintaining access

Maintaining access to the targeted system can involve recovering passwords from data that has been stored in or transmitted by the computer system (a process called password cracking) (Conrad, et al., 2010). This can involve spoofing and session hijacking. Spoofing is where the attacker pretends to be another user to gain access. The real user plays no role in spoofing which means that spoofing only requires the attacker and the machine. Whereas, in hijacking the attacker takes full control over the user's existing session therefore involving an authenticated user, the machine and the attacker (Cole, 2001). Cleaning system logs is one of the final steps of penetration testing. This endeavour enables the hacker to remain anonymous by removing all traces of their activity (Stiawan, et al., 2017). By using these hacking tools and techniques in a professional manner, a red team can help an organisation determine potential security threats such as unpatched vulnerabilities or unusual activity on the network and if a critical issue is identified early on, this is typically flagged immediately to the business so that it can be fixed (Nicholson, 2019).

Reporting findings

At the end of the red team engagement, the blue team provides the red team with any Indicators of Compromise (IoCs) that were detected during testing. The red team analyse the data collected by the blue team and incorporate it into the final report that they supply to the organisation. Such an exercise requires cooperation between both teams enabling the blue team to explain their Tactics, Techniques and Procedures (TTPs) while it simultaneously affords the red team the opportunity to offer advice to the blue team on how to detect and respond better to offensive methods in future incidents.

4.3 Ethical analysis

Despite red teaming being considered the gold standard to find vulnerabilities in organisations (Gallagher, 2020; Wright, 2018; Sloane Risk Group, 2020; Horne Cyber, 2020; Howard, 2017)) and its growth within the cybersecurity community, my academic searches[27] surprisingly failed to discover ethical analyses that examine the intricacies of red teaming and the potential impact a red team engagement may have on key stakeholders in an organisation. It is particularly surprising as red teaming is a practice that involves deception and manipulation of key stakeholders in the absence of consent: it is an unregulated, time sensitive practice and the results are limited to the knowledge of those conducting the tests and known vulnerabilities at the time of testing. I believe it is thus a necessary endeavour to examine the ethics of red teaming.

The case in consideration is whether it is ethically appropriate for an organisation to employ a red team for the purposes of finding security vulnerabilities to fix. Red teams simulate adversarial attacks using various tools to reveal weaknesses in people and technology. The aim is to not cause harm, but rather determine how vulnerable the organisation is to a real cybersecurity attack. The sponsor organisation has an internal blue team who are not made aware of the red team testing. The employees of the organisation are also unaware of the testing.

[27] The author conducted searches on databases Scopus and Web of Science using search string, ‘red’, ‘team’, ‘cybersecurity’, ‘ethic’. Both searches yielded 0 results on 11 November 2020. On the same day, the search string was widened to ‘cybersecurity’, ‘red’ and ‘team’, in the Web of Science database which yielded 21 results. 0 resources mentioned the ethics of employing a red team. The author applied the same search string to the Scopus database which yielded 30 results. 21 of those 30 were the same resources found in the Web of Science database. From the remaining 9 results, 1 resource is relevant and included in the body of this chapter (Pienta, et al., 2018). The author applied search string ‘social, ‘engineering’ ‘ethic’ ‘cyber’ ‘security’ to Scopus and Web of Science on 18 November 2020 which yielded 6 results from Scopus including a relevant paper by Hatfield (2019) on social engineering undertaken in PEN testing from which the author found Mouton, et al., (2013). This search string yielded 3 results from Web of Science, none of which were relevant.

4.3.1 Normative stakeholders' interests

As previously explained in Chapter 3, there is no specific set list of normative stakeholders as this allocation depends on the organisation's mutually beneficial cooperative scheme and the decision to be made. As we are already aware, the case presented in this chapter is based on whether it is ethically appropriate to employ a red team to find security vulnerabilities. In respect of the mutually beneficial cooperative scheme, Phillips explains that a stakeholder is party to an organisation's mutually beneficial cooperative scheme when there is the voluntarily acceptance of benefits which involves sacrifice and contribution on the part of the stakeholder and the organisation. He states that

> "[b]y doing so, the organisation has incurred obligations of fairness to attend to the well-being of these stakeholders – at least insofar as their well-being is affected by interactions with the focal organisation. This will typically include groups such as financiers, employees, customers, suppliers, and the local communities" (Phillips, 2003b: 217).

Based on the decision to be made and the assumption that the following groups are in a mutually beneficial cooperative scheme with the organisation, shareholders, employees, customers, suppliers and the local community are listed as normative stakeholders. As the blue and red teams play a pivotal role in this scenario, both teams' interests are respectfully considered under the employees' interests and suppliers' interests.

Shareholders' interests

Shareholders have an invested interest in increasing the value of their shares (also known as stocks) in the sponsor organisation. When shareholders sell their stocks at a higher price than the original purchase they make a profit known as a capital gain. At the most fundamental level, stock prices change due to supply and demand. Meaning, if more people want to buy a stock (demand) than sell it (supply), the price moves up. Conversely, if more people want to sell a stock than buy it, there would be a greater supply than demand and the price would fall. Bianchi and Tosun's (2019) research indicates that cybersecurity breaches have a strong negative impact on firms' stock returns. Cavusoglu, et al., (2004) found that announcements of security breaches negatively affect stock prices, with the most significant effect occurring two days after the public announcement of the breach.

Shareholders have interests in the value[28] of the organisation in which they have invested. A company's value is typically affected by its earnings, i.e. the profit a company makes. It is up to the board of directors to forecast long term earnings and determine whether there is enough to pay for ongoing and future business activities (called retained earnings) as well as to divvy out a portion of

[28] The value of a company can be determined by multiplying the share price by the number of shares not sold.

earnings to shareholders on a monthly, quarterly, biannual or annual basis. Akey, et al. (2018) show that corporate data breaches impact firms' value in the long term. While Bianchi & Tosun (2019) found that a breach can influence a firm's reputation by drawing negative attention from investors, consumers and the general public. They claim that "the risk of being hacked effectively represents a reputational risk, with consequences that span from increasing operating, capital or regulatory costs, to the destruction of shareholder value, consequent to a reduction in firms' profitability and revenues" (Bianchi & Tosun, 2019: 26). If we take it to be true that breaches negatively impact stock prices and company value, it is fair to assume that shareholders have a legitimate financial interest in investing in a company that is cognizant of the threat landscape and takes adequate action to reduce the risk of falling victim to a cybersecurity attack.

Red teaming is one of the most advanced ways to manage the two main security weaknesses in organisations, i.e. people and technology. If a red team engagement can discover and remediate known vulnerabilities in people and technology, red teaming can consequently reduce the threat landscape, suggesting that employing a red team is in the interests of shareholders. However, employing a red team who introduce or keep any *backdoors*[29] open in the system or discuss the results of the simulated attacks with unauthorised persons, would not be in the interests of shareholders. Such actions would introduce new security risks to the organisation. If those risks lead to a security breach, it could jeopardise the firms stock returns and value.

It is in shareholders' interests for red teaming to model real world conditions insofar as possible because this will produce a true snapshot of the security posture of the organisation. In order to model real world conditions, the testing needs to include the deception and manipulation of employees (including the blue team) as this is often a tactic used by adversaries to successfully target enterprises (Resnik & Finn, 2017). If deception and manipulation are necessary to model a real attack, employees cannot be forewarned of the testing as this may alter their behaviour. In addition, they cannot be forewarned as this would not reflect a real-life attack because a real adversary does not give employees prior notice of their intentions to attack, nor do they request consent to launch an attack.

While it is in shareholders' interests in profit, growth and expansion to employ a red team to conduct necessary tests required to mimic a real attack, it is equally important to shareholders that the engagement is limited in scope, causing little to no unnecessary interruption to daily functions. If the red team engagement is unlimited, this can negatively affect the organisations ability to execute daily

[29] A backdoor is a method used to bypass normal security measures and gain high level user access (known as root access) on a computer system, network, or software application. Backdoor are gateways that can be used to steal personal and financial data, install additional malware, and hijack devices (Malwarebytes, 2020).

functions, which may affect shareholders' interests in profit and growth. For example, red team testing can take up to 12 weeks to complete. Certain tests can put strain on a business and bring devices down, while scanning and exploitation attempts can increase the overall attack surface (Oakley, 2018). This can be problematic and introduce large delays, excessive losses and service interruptions (Mirkovic, et al., 2006).

It is equally important to shareholders that the results of the engagement are clearly communicated to the relevant persons such as the blue team and management as this will contribute towards satisfying the purpose of employing the red team, which is to find vulnerabilities to fix.

Employees' interests

Employees have an interest in keeping their own personal information stored by their employer private and confidential, e.g., banking information, salary, home address, tax number. If this sensitive information is breached in a cybersecurity attack, Brey (2007) argues that this can have a negative psychological impact on data owners. If a red team engagement increases the security of the data held by the organisation, it can place the organisation in a better position to protect employees from potential harm. Therefore, employing a red team could be argued as being in the interests of employees' health and safety.

The average cost of a data breach can be significant – estimated to be 148 USD for every compromised employee or customer record (Ponemon Institute & IBM, 2018). The financial cost of a data breach can greatly impact employees' interests. For example, plummeting stock prices and a reduction in gross earnings can affect all employees as it can reduce the organisation's financial flexibility to offer employees competitive pay and perks in the short and long term (Bianchi & Tosun, 2019). Employees who own company stocks, in other words employees who are also shareholders can suffer even more from a security breach due to a decline in stock prices and company value. It is thus in employees' financial interests for management to implement measures that will reduce the likelihood of falling victim to a cybersecurity attack. Red teams can help protect those interests by identifying vulnerabilities in people and technology that adversaries would otherwise exploit.

Employees pose as a significant security threat to enterprises due to organisations' adoption of new digitised working practices, e.g., sharing multimedia content including audio and video files and sending emails. Adversaries target employees via social engineering attacks to create a gateway of access by encouraging employees to download a malicious attachment, go onto an insecure website, running unsafe software or by sending confidential information to an impersonator. In fact, social engineering attacks represent the lion's share of access vectors for cybersecurity incidents (Oakley, 2018). Adversaries are indifferent as to the lines they will cross in order to hack an organisation and red teams will replicate adversaries' actions, to a reasonable degree. For example, it is common for red teams to impersonate a stranger and set up social media account. Social media profiles are typically a rich source of data that can be used to learn who to target in order to gain access to company facilities, systems or confidential information. Red teamers also use physical intrusion testers to solicit critical information from employees under false pretences (DeMarco, 2018). For example, a red teamer may physically impersonate a package delivery person to gain physical access to the sponsor organisation. Or they may obtain a cloned building access card to get access to a server room. Once inside the organisation, the red teamer can initiate a successful hack by accessing an unattended computer or by scattering infected USB drives around an open office in the hope that one employee plugs the infected device into one computer (DeMarco, 2018). While these tactics directly target employees, it is reasonable to suggest that the tactics adopted are simply used to reflect those exercised by real adversaries and to measure how susceptible the sponsor organisation is to a social engineering attack (Secarma, 2020). Accordingly, it is fair to assume that identifying all known security weaknesses in the sponsor organisation is in the security interests of employees due to the security gains to be made from identifying how an attacker may gain access to the organisation.

However, if a red team utilises social engineering tactics that are limitless, highly intrusive and unnecessarily manipulative to coerce employee behaviour, in a private or professional capacity, this is not in the interests of employees as it may cause harm. DeMarco (2018) states that red teams should not be authorised to use employees' credentials to access the company's systems. In cases where organisations authorise the red team to leverage employees' credentials to determine the extent of a threat and impact, this should only be conducted in limited and secure situations. Otherwise, the fact that the credentials were obtained should be noted, but the security team should proceed no further (DeMarco, 2018).

Employees are an additional threat to the sponsor organisation as they contribute towards accidental human-centric data breaches which can be initiated by sending information to the wrong recipient, sharing information without permission, or not knowing security protocols (Redscan, 2020). Mistakes of this nature are considered an operational reality for modern organisations and may be lessened through cybersecurity awareness, education and training (Yaghmaei, et al., 2017). To circumvent the potentiality of accidental errors, it is in employees' personal and professional interests to learn about good cyber hygiene and learn how certain practices can compromise the security of an organisation. A red teaming exercise offers an organisation the opportunity to identify human-centric security weaknesses in organisations which can include flagging accidental errors executed by employees.

Following-up a red team exercise with appropriate cybersecurity education and training is the interests of employees' professional development.

There are reported instances in organisations where employees are not supplied with the tools needed to share sensitive data securely. For example, 55% of employees intentionally shared data in a manner that broke company rules because their organisation did not have the secure tools needed to share sensitive data (Opinion Matters, 2019). Organisations who fail to provide the necessary tools employees require to safely store, process and share information with colleagues and clients are working against the interests of their employees as well as other stakeholders e.g., shareholders, customers and the local community. It is unreasonable to force employees to choose between 1) getting the job done and risking their own security, the security of the organisation as well as the data being shared and 2) not getting the job done and risk being reprimanded for not completing the task at hand. It is possible for an employer to not know that the software they use is insecure. It is equally possible for management to know about an insecure tool but for one reason or another, e.g., lack of finances or ignorance, the problem was not remedied. An external red team can identify insecure practices like these which would be in the interests of all employees.

Data breaches can be instigated by malicious employees who intentionally put the security of the organisation at risk. A 2019 Insider Data Breach survey reveals that 32% of employees would consider taking company information to a new job, 23% actually stole company data and brought it to their new job and 13% shared or removed data because they were 'upset at the organisation' (Opinion Matters, 2019: 7). Stealing sensitive data, misusing access to networks, applications and databases to cause damage or disruption and/or erasing or modifying sensitive data are just a handful of the ways that malicious insiders can cause harm to an organisation and fellow employees (Redscan, 2020). Organisations try to manage malicious insider threats by developing codes of ethics to guide ethical employee behaviour or by using general deterrence techniques to discourage computer or data misuse (Yaghmaei, et al., 2017). Organisations can also employ a red team to protect themselves from malicious insider attacks via insider threat simulation. Insider threat simulations identify weaknesses that stem from inside the firm that may be exploited. Given that innocent employees can negatively be affected by a malicious colleague who steals or misuses private and confidential data, it is reasonable to assume that employing a red team is in the interests of non-malicious employees and conversely, in conflict with the interests of malicious employees.

Employees' perception of red teaming is completely dependent on the perception of the organisation and the goals of the red teaming exercise. For example, surreptitiously employing a red team to target employees in the absence of consent may be perceived by employees as an infringement of distributive justice or as an ill-willed attempt to 'catch them out'. Providing supportive cybersecurity training to employees as a follow-up to the red team engagement may circumvent the potential issue of red teaming being viewed as a form of distributive injustice. For example, giving employees clarity on the intention of the engagement, the vulnerabilities found and how those vulnerabilities can be better managed in the future may provide employees with transparency as to the purpose of the engagement and it may also remove the potential for misunderstanding or confusion all the while fortifying the organisation's resiliency to an attack. All of which advance employees' interests. In the absence of clear communication and follow-up training, it can be easy for employees to 'think the worst' and view the exercise as a way for management to pinpoint employee weaknesses and reprimand them for their incompetence.

As regards conducting red teaming in the absence of consent, the engagement could be explained to employees as a necessary form of security research that cannot be conducted if consent is obtained prior to testing. This would require employees to have a prior understanding that the primary benefit of red teaming is that it mimics real-world attacks. A real-world attack cannot be simulated to the fullest degree if one of the main vulnerabilities in the sponsor organisation, i.e. employees, are informed of the planned attack. Obtaining informed consent prior to testing can reduce the usefulness of red teaming if the informed individuals adjust their behaviour (Hatfield, 2019). If employees alter their behaviour to align more stringently with good security practices and protocols, the results of the simulated attack will not be a true reflection of the security behaviours existent in the organisation. Unless employees are made aware as a matter of course that red team testing may be done at any time, even as a deterrent to bad behaviour. In fact, it is likely that the results of the testing will describe the security of the sponsor organisation in a more favourable light which could create a false sense of security. If we assume that employees understand that red teaming is a form of security research that cannot be truly executed if employee consent is obtained prior to testing, it is possible that employees will not perceive red teaming as an injustice.

There is no regulatory body that compels ethical hackers like red teams to undergo ethical training. Hatfield (2019) declares that security companies that engage in human hacking such as social engineering, have an obligation to include ethical training "as part of the development and employment of white hat social engineers" (Hatfield, 2019: 343). According to Harrington (1996), an absence of an industry wide professional code of conduct makes it easier for individuals to rationalize irresponsible behaviour as codes of ethics are a means of deterring unethical behaviour and can be a basis for internal sanctions (Harrington 1996). It would be in employees' interests for the sponsor organisation i.e., their employer, to ensure that the persons conducting the security testing have undergone some form of ethical training and/or take ethical underpinning from accredited bodies like the ACM.[30] The first two obligations listed in the ACM Code of Ethics and Professional Conduct relate to contributing towards society and human well-being, and avoiding harm (ACM, 2018).[31] It states that in situations where well-intended actions can result in harm, those responsible are obliged to undo or mitigate the harm as much as possible. If red teams are privy to an ethical code like that prescribed by the ACM, red teams may appreciate the impact that their engagement may have on stakeholders before testing commences. This could help mitigate any potential harm caused by ruling out actions that may cause harm before testing commences. For example, gaining physical access to a sensitive room by breaking into the home of one of the employees who has a key to the room would not be appropriate and obviously be adverse to the interests of employees.

Hatfield (2019) argues that certain tactics such as phishing can cause psychological harm to the targeted individuals. A study surveying approximately 500 persons working in security and non-security positions was presented at ShmooCon 2020 in Washington DC which reveals that employees working in legal, human resources or at a reception desk are nine-times more likely to object to receiving a phishing email as part of a red team engagement when compared to a security professional such as a red teamer or incident responder (Whittaker, 2020). The same study found that those who conduct the simulated attacks are also likely to object to specific tactics being used against them personally e.g., phishing emails and planting compromising documents. The motivations behind the objections is not clarified and there could be multiple reasons for the objection. For example, the objections may stem from a fear that a contribution to the success of a simulated attack may result in job loss. Or the objections might be based on the understanding that certain tactics like phishing should only be conducted once consent has been obtained. Or, perhaps these tactics are perceived as causing more harm than good. Regardless the grounding for the objection, this survey suggests that certain red teaming tactics are not in the interests of employees or the blue team.

In terms of reporting the results of the engagement to the organisation, it is in the interest of employees that management employ a red team who will provide a report that includes a clear explanation of the technical and human-centric weaknesses that contributed towards the success of the simulated attack. This includes naming employees who pose a security threat to the organisation through either their lack of cybersecurity competence, trusting nature or otherwise. The purpose of identifying such employees is to improve the overall security of the organisation through better

[30] The ACM Code of Ethics and Professional Conduct lists computing professionals' ethical responsibilities and is relevant to ethical hackers, red teams and information security experts because it is designed for all computing professionals "including current and aspiring practitioners, instructors, students, influencers, and anyone who uses computing technology in an impactful way" (ACM, 2018).

[31] Harm is understood as negative consequences that are significant and unjust i.e., physical or mental injury, unjustified damage to property, reputation and the environment (ACM, 2018).

training and education. In such cases, identifying employees who contributed towards the success of the simulated attack is in the interests of employees.

It is obviously not in the interests of employees if the sponsor organisation requests that employees names are included in the red team report for the sole purpose of dismissing or reprimanding the employee. Mouton, et al., (2013) state that is it ethical to report the names of the employees to the sponsor organisation despite the potential negative consequences for the employee (Mouton, et al., 2013: 6). Yet, they do not define what they mean by 'potential negative consequences'. In the author's professional experience, naming and shaming those who contributed towards the success of simulated attacks can have serious negative repercussions for employees. For example, under Chatham House Rules, an ethical hacker shared a client case with the CANVAS Project consortium at a CANVAS Cybersecurity Workshop (2018).[32] The red teamer explained that once testing was complete, the red team submitted a report to their client which included the name of the employee who was successfully targeted and exploited during the simulated attack. Weeks after the report was submitted to the client, the red teamer discovered that the employee in question had been dismissed.

The true purpose of a red team is to find ways to improve the blue team by testing the defences of the organisation (Miessler, 2016). To increase the realism of a simulated attack, the Chief Information Security Officer (CISO) and the head of physical security are sometimes deliberately excluded from the details of the engagement (DeMarco, 2018). This is the case in Black Box[33] testing where the blue team have no knowledge of future red team testing *and* the red team are given no information relating to the organisations defences, applications etc prior to the start of testing (DeMarco, 2018). Those privy to the details of the red team test in Black Box testing are typically limited to a handful of senior management members. Blue teams will be cognizant of the fact that there is a lack of ethical standardisation in the field of offensive cybersecurity. For example, in terms of the rules of engagement, red teams can follow the Open Source Security Testing Methodology Manual (OSSTMM) which provides a methodology for analysis and measurement of operational security (Herzog, 2010). The purpose of the manual is to improve operational security - it does not include any reference to ethical behaviour. Blue teams may also know about the TIBER-EU framework designed in 2018 to standardise the way red teams perform intelligence-led tests across Europe (European Central Bank, 2018a; European Central Bank, 2018b). The framework encourages red teams to follow rigorous and

[32] Cybersecurity in Business. Helsinki, Finland. 2018. Closed door event involving legal, ethical, technical experts in Information Security. Invitation Only. May 2018.

[33] There are three types of testing. We focus on Black Box testing as it is the most likely out of the three types of testing to raise ethical issues. Grey Box involves the organisation supplying the red team with some information relating to the target systems and the blue team are given notice of the test. White Box testing is when the red team are given full details of the network, applications and internal procedures and the blue team know the full details of the test in advance.

ethical red team testing methodology (European Central Bank, 2018a: 25) and practice strong ethical behaviour – both of which are undefined (European Central Bank, 2018a:40). In the absence of ethical standardisation, the blue team have an invested interest in management employing a red team who have the ability to rule out inappropriate or unnecessary testing or actions before testing commences as this can avoid unnecessary harm being caused to the organisation, systems and people that the blue team endeavour to defend.

It is in the interests of the blue team for management to choose a trustworthy red team. This could be a red team that has undergone ethical training, follows a code of ethics or at the very least, are members of an industry-recognised security organisation that compels members to follow strict ethical guidelines during testing and reporting. For example, red teams can voluntarily become a member of an industry recognised accredited organisation like CREST.[34]

It is also in the blue team's interests for management to restrict red team tactics to a degree that enables the red team to simulate real-world attacks but does not cause unnecessary stress or strain on the organisation's people and technology. For example, if a red team is granted permission to use extraordinary measures that are limitless with no known bounds, this could result in property damage or a physical injury to the organisation, its systems and applications and its people. Examples of using unnecessary tactics and causing harm include initiating a DDoS attack that significantly impacts services. Similarly, if management restrict red team testing to the extent that it nullifies the effectiveness of a red team exercise, this is not in the blue team's interests. If the exercise does not mimic a real-world attack due to constraints placed by management, the red team's purpose of mimicking an attack is an impossibility. In addition, if restrictions negate the purpose of red teaming, the results of the engagement will not be a true reflection of the security posture of the organisation. Skewed results can give the organisation and the blue team a false sense of security which is not in the blue team's interests.

While it is not mandatory for red teams to offer expert advice on counter measures that can be adopted to better manage the vulnerabilities identified, some red teams do. This can be worthwhile as the blue team can substantially gain from acquiring information about how to improve or better manage the discovered vulnerabilities. Such information can also guide future cybersecurity Vulnerability Management Programs (VMPs).

[34] CREST is a security organisation that has developed their own code of ethics in an effort to influence the judgment of professionals working in the security industry (CREST, 2019). CREST also provide certifications like the EC-Council (Thomas, et al., 2018) and have their own list of accredited security companies.

Circumstances can arise where the red team do not have the ability to clearly explain, interpret or contextualise the technicalities of their results to the blue team in a simple and understandable way. This can create problems for the blue team if they do not understand what they need to do in order to fix the identified vulnerabilities. It is not in the blue team's interests for the sponsor organisation to engage with a red team who is either unable or unwilling to 'dumb-down' the results of the simulation to the blue team. This can happen in cases where the red team have an elitist[35] point of view (Miessler, 2016). Similar to employees, it is not in the interests of the blue team for management to use the results of a successfully executed simulated attack as a cause for dismissing members of the blue team.

Customers' interests

Customers have a financial interest in organisations that provide convenient access to services and competitive prices. Financial interests can be affected if an organisation falls victim to a cybersecurity attack and is forced to redistribute resources in order to cope with the associated costs. Cybersecurity expenses continue to grow in proportion to the increasing frequency of cybersecurity attacks on enterprises. In order to compensate for rising security costs, organisations may need to increase the price of products or services, which is not in the interests of customers who desire value for money and competitive prices. If an organisation can lower their risk of being attacked, this can reduce the cost of cybersecurity insurance. One way to lower cybersecurity risk is to employ a red team to identify vulnerabilities that need to be fixed. Therefore, employing a red team could be argued as advancing customers interests.

In Europe, EU citizens place high value on safety and privacy in cyberspace (Wenger & et al, 2017). As some cybersecurity breaches involve the misuse, distribution, and manipulation of customer data, protecting customer data from adversaries is particularly important to customers who consider their personal information to be private and confidential. Employing a red team can satisfy these customers interests because it reinforces an organisation's defences. It goes without saying that it is in customers' interests for red team testing to be undertaken in a manner that does not compromise, misuse or modify private and confidential customer information.

Suppliers' interests

Suppliers have similar interests to shareholders and customers. They have an economic interest in the frequency of attacks declining as an increase in attacks can influence cybersecurity costs. Suppliers such as CSPs have an interest in their cyber environment remaining protected at all times from

[35] Red teams are considered the highly esteemed experts in cybersecurity when it comes to finding security vulnerabilities. As a result of power and position they hold, it is possible for red teams to develop an elitist attitude where they view themselves better than regular security professionals.

adversaries. Suppliers who heavily rely upon cyberspace to execute their daily functions have an economic interest in partnering with businesses who appreciate the importance of protecting and securing interconnected systems, data and people. Employing red teams can thus be argued as being in the interests of suppliers. More specifically, employing a red team is in the interest of suppliers when the red team engagement does not harm the service supplied to third parties.

The red teams, as a supplier, have a financial interest in finding vulnerabilities in cyberspace as it is their main source of income. In order to find vulnerabilities, vulnerabilities must exist. Without the continuous existence and/or emergence of new security vulnerabilities in organisations, red teams do not have a function. This means that red teams have an interest in proving to the sponsor organisation that the costly work they carry out provides a higher level of protection over and above other 'standard' methods of protection, i.e. passive and active defence. This means that red teams have a vested interest in utilising various approaches and tools that are likely to provide them with access to the sponsor organisation. Using measures adopted by real adversaries allows the red team to simulate a real-life attack. Even though such measures include manipulating both people and technology, it is in red teams financial and reputational interest to do so. If red teams can demonstrate their 'worth' to the sponsor organisation by finding security vulnerabilities, it is likely that management will perceive the exercise as justifiable, valuable, necessary and worthwhile.

It is equally in the red teams' interest to be perceived as trustworthy by business organisations. Given the unprecedented access that red teams can acquire during an engagement, it is vitally important to red teams (and the organisation) that in the event of gaining access to trade secrets or other sensitive or valuable information that the information remains private and confidential. Trust between the organisation and the red team is the thread that sows the initial contract together and secures future employment (most red teams are sourced through word of mouth). An abuse of power would negatively impact the red team and their reputation and thus not be in their interests. In addition, unnecessarily causing harm to those who are being manipulated could also be viewed as an abuse of power by the red team and negatively impact future relations between them and the organisation. Therefore, it is in red teams' interest to establish a clear and concise document stating the Rules of Engagement which includes the scope of manipulative measures used to gain access to the organisation as well as the protocol to follow should the red team come across any sensitive or valuable information. It is also in the red teams interests to demonstrate that they followed agreed protocols to avoid any accusations of misdemeanour. Consequently, clear communication between the red team and the organisation before the engagement begins is critically important to red teams as it can help clarify both parties expectations and iron out any anticipated conflicts of interests.

Local community interests

Generally, employees reside within close proximity to their workplace. Thus, the interests of employees and the local community are closely connected. Like employees and shareholders, local communities have long-term economic interests in organisations in their region minimising the cost and collateral damage caused by cybersecurity attacks. If employing a red team increases a business's ability to detect, manage and respond to potential cybersecurity breaches, this will satisfy the interests of the local community. Organisations who utilise advanced cybersecurity measures to protect their assets are also protecting their employees', shareholders' and suppliers' interests – all of whom may be members of the local community.

4.3.2 Derivative stakeholders' interests

Derivative stakeholders are those who can help or harm the organisation or the interests of the normative stakeholders. Phillips (2003b: 222) states that it is in the interests of the organisation for management to spend a limited amount of time and resources attending to derivative stakeholders' interests because doing so is in the interests of the organisation and its normative stakeholders. In respect of the decision to employ a red team to find vulnerabilities, competitors are listed as the first derivative stakeholder as it is not uncommon for organisations to believe that attacks stem from competitors with the aim of sneaking an advantage in the marketplace or to steal private information (Kaspersky, 2015). If competitors pose as a security threat, they require the attention of management. Similarly, the media can report red teaming in a positive or negative light which may influence and alter the opinions of readers, who may include shareholders, employees, customers, or suppliers. For example, if red teaming is described as a slight against employees by the media, it may negatively affect employees' attitude or behaviour. The media are thus listed as the second derivative stakeholder.

Competitors' interests

Competitors who are interested in keeping the costs of cybersecurity down in their sector have a financial interest in reducing the prevalence of successful attacks within their sector. If employing a red team reduces the prevalence of attacks in one sector, this is in the interests of like-minded competitors. Equally, competitors have an interest in the security products, policies, practices and processes undertaken by their rivals. In knowing the practices of their rivals, competitors can benchmark their own processes and strategies and implement similar or better strategies to 'get ahead'. Competitors can benefit from their rivals adopting a reasonable standard of cybersecurity. For example, competitors can experience a change in trading conditions or may face changes in the perception of their industry as a result of the behaviour of their industry counterparts. It is thus, not in competitors' interests for their rivals to be ignorant to good cybersecurity policies and practices.

Competitors have an equal interest in outdoing their rivals. This extends to competitors aspiring to have better cybersecurity practices and policies than their counterparts as this could place them in a better position to withstand and respond to cybersecurity attacks. It could thus be argued that competitors have an interest in employing red teams themselves to reduce their own attack surface. This is particularly relevant in cases where adversaries launch several attacks on a specific industry and the competitor gains from their rivals being successful targeted resulting in more market share for the competitor.

As competitors can profit from their rival being successfully attacked, it is not uncommon for targeted organisations to blame competitors for launching cybersecurity attacks (Kaspersky, 2015). If this is true and competitors are launching their own cyber security attacks to 'get-ahead', it could be argued that competitors have an interest in the targeted organisation not employing a red team to fortify defences against attacks as this would reduce the attack surface for the competitor to launch targeted attacks.

It should be noted that it is not in the interests of organisations within one sector for it to become routine practice to launch attacks on one another. This is breeding ground for distrust and possible cases of defamation.

News Medias' interests

The decision to be made is whether it is ethically appropriate for an organisation to employ a red team to find vulnerabilities. The news media can help or harm an organisations' decision based on the context of the story they publish. For example, the media can write about an organisation that employed a red team to pinpoint human-centric vulnerabilities with the intention of dismissing employee(s) who contributed towards the success of simulated attacks.

The news media will have an interest in an organisation who employs a red team who, for example, purposefully leak information about the deplorable security practices of a large multinational corporation. 'Deplorable security practices' could be interpreted as anything from the distribution of malware to potential customers as a source of revenue, to employees' misuse and abuse of office computers. Both would create negative publicity for the sponsor organisation.

The media may also have an interest in an organisation who employed a red team who did not uncover any vulnerabilities but later suffered from a detrimental cybersecurity breach. Sensationalist news stories such as these create hype and interest, and it is in the interests of the news media to cover them regardless of the whether the story will help or harm the organisation and its normative stakeholders. It is thus in the organisations' interests to allocate time and resources to effectively manage the news media appropriately.

4.3.3 Prioritising conflicting stakeholders' interests

Deception & harm

The results of the stakeholder analysis suggest that normative stakeholders significantly gain from an organisation employing a red team to find vulnerabilities in people and technology. These stakeholders include employees who are not informed of the testing before it begins and are deceived by red teams in order to determine how susceptible they are to social engineering attacks. If red teams were to obtain consent from employees prior to testing, it would defeat the entire purpose of a red team test, i.e. to find real security vulnerabilities in people and technology, and would be a misrepresentation of a real-world attack as adversaries do not obtain consent from their target before executing their attack. Therefore, some level of deception is necessary for a red team to achieve its objective. Yet, harms can be caused to normative stakeholders (not just employees) in the short and long-term if 1) the engagement is not limited and unnecessary intrusive methods are used against employees and technology to execute simulated attacks, 2) constraints in the rules of engagement are too strict preventing real-world attack simulation and creating a false sense of security, and 3) reporting is used as a cause for dismissal or disciplinary action.

It is in stakeholders' interests for these harms to be mitigated and the analysis suggests that they can be mitigated if 1) the red team engagement is limited to the extent that it simulates real-world attacks in a fashion that does not cause unnecessary harm to people or technology, 2) the constraints of the rules of engagement do not negate the purpose of a red team engagement i.e. to find vulnerabilities in people and technology, 3) red team reporting includes the names of employees who contributed towards the success of the simulated attack for education and training purposes only, and 4) the red team have the ability to and are willing to constructively share the technicalities of the results with the blue team.

Limiting the engagement

It is unequivocally in the interests of shareholders, customers and suppliers for the organisation to mimic cybersecurity attacks that include deception as it improves the security of the organisation. Yet deceiving employees and manipulating technology raises a conflict of interest for employees and the blue team. Phillips (2003a; 37) states that it is possible to resolve conflicts of interests conceptually by taking into account whether the proposed action supports the continuation of the cooperative scheme and is likely to reach the communicative assent of all stakeholders. A red team engagement that is unlimited and intrusive may cause harm to employees and technology which conflicts with employees and blue teams' interests. Aggressive tactics could be argued as supporting the continuation of the cooperative scheme but it is unlikely to reach the communicative assent of all stakeholders. In contrast, constraining a red team engagement insofar as it prevents the red team from mimicking a real-world attack on the sponsor organisation, i.e. obtaining consent from employees before the engagement and objecting to social engineering attacks, defeats the whole purpose of a red teaming exercise. Restricting red teaming in this way can create a false sense of security which conflicts with all normative stakeholders interests in improving the security of the organisation. It is difficult to imagine how it could be argued that spending money on an exercise that provides the organisation with a false sense of security and fails to satisfy its purpose supports the continuation of the cooperative scheme. It is also difficult to conceive that employees will agree to employing a red team to conduct limited testing that does not satisfy its purpose.

Industry accredited red teams

Red teams can engage in unethical behaviour if they have the means, opportunity and motive to do so (Pendse, 2011). Consequently, choosing a reputable red team who will remain compliant with the rules of engagement can be crucial. There is also the predicament that the results yielded from a red team engagement hinges on the expertise of the red team and the known vulnerabilities at the time of testing. Finding an industry-accredited red team is perhaps one way the sponsor organisation can circumvent the lack of ethical governance and standardisation in this field. Choosing a red team who abide by an ethical code of conduct and follow some form of standardisation would not only be in the interests of all normative stakeholders, but would also support the continuation of the cooperative scheme and be likely to reach the assent of all normative stakeholders.

Rules of engagement

In terms of deciding upon the rules of engagement, it is in all stakeholders interests for management to treat the process with a high level of scrutiny. This includes conducting a background research on the red team as well as researching the proposed testing methods. It is necessary for management to engage in open dialogue with the red team before testing commences. Open dialogue will give the red team the opportunity to contextualise the associated risks of testing as well as afford the red team the chance to respond to the organisation's request for testing. The red team can also clarify what is possible or impossible, what the security risks are, and how testing may impact the organisation and its stakeholders. The red team can outline what they expect from the organisation once the testing is complete, e.g., that the organisation will remediate any discovered vulnerabilities, where possible. Open dialogue will afford management the opportunity to assimilate the information provided and based on the level of risk introduced by a test to their people and their systems, management can determine what is acceptable and what is not. The red team and management can thereafter draw up the rules of engagement which should include how the results will be reported to the organisation. This proactive approach could provide management with the information they need in order to make a decision that supports the continuation of the cooperative scheme and is likely to achieve the communicative assent of all stakeholders.

Reports that improve security

Including the names of employees who pose as a security threat to the organisation affords the employee(s) and management the opportunity to further educate and improve awareness of good cyber hygiene. This is in the interests of all normative stakeholders as it contributes towards improving the security posture of the organisation. In contrast, using a red team report as a cause for dismissing people within the organisation who were subjected to a simulated social engineering attacks, is not in the interests of employees. While pruning persons who posed as a security threat to the organisation may support the continuation of the cooperative scheme, it is very unlikely that employees will agree to such tactics. Based on this assumption, it is unlikely that this type of action will result in assent from all stakeholders.

Once red team testing is complete, ordinarily the red team submits a report to the organisation. It is in the interests of all normative stakeholders for the red team to constructively share the technicalities of the results with the blue team as this will compound the benefits of the engagement. Ensuring that the employed red team has the ability and are willing to engage in open dialogue with the blue team post-assessment not only supports the continuation of the cooperative scheme, it is also likely to reach the communicative assent of all stakeholders.

4.4 Conclusion

The normative analysis in this chapter suggests that it is ethically appropriate for organisations to employ red teams to find vulnerabilities in people and technology when 1) the engagement is limited to the extent that it simulates real-world attacks in a fashion that does not cause unnecessary harm to people or technology, 2) the constraints of the rules of engagement do not negate the purpose of a red team engagement i.e. to find vulnerabilities in people and technology, 3) red team reporting includes the names of employees who contributed towards the success of the simulated attack for education and training purposes only, and 4) the red team have the ability to and are willing to constructively share the technicalities of the results with the blue team. When red teaming satisfies these four conditions, it is in the interests of all normative stakeholders to employ a red team to find security vulnerabilities in people and technology. Such action supports the continuation of the cooperative scheme and is likely to reach the communicative assent of all stakeholders.

Chapter 5 Responding to Ransomware Attacks

5.1 Introduction

This chapter focuses on one common threat facing organisations today: ransomware attacks. A ransomware attack generally involves encrypting valuable data or blocking access to devices for ransom. It is proposed that organisations can respond to ransomware attacks in one of three ways; by paying the ransom, negotiating a lower fee, or by not paying at all. An ethical analysis is conducted to establish which out of the three available responses is the ethically appropriate response. Phillips' systematic conceptual method is adopted to examine the ethics of responding to ransomware attacks based on Robert Phillips' work on stakeholder theory and obligations of fairness (Phillips, 1997; Phillips, 2003a; Phillips, 2003b). At the time of writing the author was not aware of a similar analysis conducted in the ethical literature.[36]

Although there is no academic literature specifically on the business ethics of paying a ransom in a ransomware attack, the generation of revenue via threat or extortion has been a long-established criminal practice and there is literature on the ethics of organisations choosing to pay a ransom in other situations. For example, Peter Singer (2014) writes an opinion piece on how governments who pay ransoms save the lives of some of their citizens but put the remainder of their citizens at greater risk. Howard (2018) provides a deontological analysis of governments paying ransoms to terrorist organisations and concludes that paying ransoms makes the state complicit in the serious injustices that ransom payments fund. Lansing & Peterson (2011) apply a utilitarian approach to determine whether commercial shipping owners should pay ransoms to Somali pirates who take the vessel and sailors hostage. They conclude that shipping owners who buy kidnapping insurance for the purpose of paying off ransoms will prove to be the industry's undoing. Whereas, not paying ransoms to Somali pirates is in the best interest of society as a whole. It is possible that corporations such as banking institutions have a policy for responding to and preventing physical kidnappings of executives for example. However, physical kidnappings are not the focus of this paper, rather responding to the digital hijacking of information and devices is. The closest analysis of ransomware attacks that touches on the ethics of ransomware that the author could find is Cartwright & Cartwright's (2019) article on ransomware and reputation. They apply economics and game theory to better understand the ransomware business model and conclude that a ransomware attacker who returns victim's access to files or devices gives criminals a good reputation from which criminals ultimately benefit.

[36] A combination of three search terms "cyber*", "ethic*", "ransom*" were used to search publications within two databases: Scopus ("article title, abstract, keyword") and Web of Science ("topic") on 22 October 2019. Both databases yielded 1 result each, the same 2016 paper "Cyber ethics and cybercrime: A deep delved study into legality, ransomware, underground web and bitcoin wallet" (Upadhyaya, 2016). A review of the full text revealed that this is a technical paper about ransomware.

Examining the ethics of responding to ransomware attacks is a relevant and necessary topic that ought to be addressed and thoroughly assessed in areas of business ethics, ethics and IT ethics as ransomware is both a growing threat to enterprises and responding to ransomware attacks can negatively impact normative stakeholders long- and short-term interests. This chapter may be particularly useful to organisations who strive to follow a stakeholder-centred approach, wish to determine whether a particular action is ethically appropriate or are faced with the dilemma of prioritising or resolving a stakeholder conflict. The conceptual method used in this chapter can be replicated by management and used as a model for establishing whether any business decision is ethically appropriate.

This chapter begins with some insights into the emergence of ransomware as a growing threat to organisations (Section 5.2.1). This is followed by a discussion on how ransomware behaves, its impact and costs on organisations (Section 5.2.2), who is a target (Section 5.2.3) and how affected organisations are choosing to respond (Section 5.2.4). Robert Phillips' work on stakeholder fairness and legitimacy is used to determine who is a legitimate stakeholder worthy of management's attention when choosing to respond to an attack (Phillips, 1997; 2003a). Phillips' general conception of ST is adopted, which is that decisions made by management must consider the interests or well-being of legitimate stakeholders. The proposed responses: pay, don't pay or negotiate, are analysed in terms of the affects each response may have on normative stakeholders' interests (Section 5.3.1). It is explored whether there is a difference between the choice of paying and negotiating (Section 5.3.2). Derivative stakeholders' interests are discussed (Section 5.3.3). To resolve conflicts of interests should they arise, Phillips' work on the prioritisation of interests to resolve same is used (Section 5.3.4). The response that advances the interests of legitimate stakeholders and supports the continuation of the cooperative scheme is considered the ethically appropriate response to a ransomware attack (Section 5.4).

5.2 Ransomware

The first ransomware attack was recorded in 1989 (Choi, et al., 2016). Its popularity did not gain traction until the mid-2000s which was around the same time as the emergence of more sophisticated algorithms such as RSA (Rivest-Shamir-Ardleman) an asymmetric encryption algorithm used by modern computers to encrypt and decrypt data (De Groot, 2019). RSA has a public encryption key and a private decryption key which allows attackers to encrypt data for ransom in exchange for the private key. Whether the attackers have purchased or developed the ransomware themselves, the malware can be delivered in a number of ways. The most common delivery system is a so-called 'phishing spam'. Phishing spam are attachments that the victim receives via email masquerading as a file they could normally trust. When the file is downloaded or opened the ransomware takes effect and installs on the victim's computer. While some sophisticated versions of ransomware have built-in social engineering tools that trick victims into allowing administrative access, more aggressive versions exploit security holes to infect computers without needing to trick users.

Once the ransomware has been installed, there are several things that it might do. Typically, ransomware is designed to "kidnap" data or devices through encrypting data for ransom (crypto ransomware) or disabling access to devices, systems or apps (locker ransomware). Other malicious variations delete data (leakware, doxware, Reveton and Tobfy) or steal data (W32.BolikA!inf; Aurangzeb, et al., 2017).[37] This chapter focuses on the most frequently executed ransomware attacks i.e., crypto ransomware and locker ransomware and refer to them collectively as typical ransomware attacks (Varonis, 2017).[38] When typical ransomware infects a targeted device, it can often remain undetected until it presents itself to the user in the form of a pop-up fee note demanding payment in return for access to the data or device. The victim is usually asked to pay the fee in the form of an untraceable Bitcoin payment as this allows the attacker's identity to remain anonymous throughout the entire process. Bitcoin is a digital currency that ransomware attackers use as it allows for anonymous financial exchanges. When ransoms are paid via Bitcoin, it can be very difficult to source the attacker. This means that the attacker does not have to concern themselves with the fear of being caught. Granting such anonymity to the attacker makes ransomware a minimal risk enterprise.

5.2.1 A growing threat & a criminal offence

What is worrying for organisations is that there is no indication that the growth of ransomware will subside. For example, as attackers can evade identification by employing untraceable payment methods via Bitcoin, ransomware is becoming a popular low risk business model. If there is no risk of getting caught, there is nothing to deter attackers' efforts. The ransomware business model provides a high reward for attackers which can also be problematic. For example, attackers have started successfully targeting organisations that cannot afford any interruption (a period known as downtime) to their services, files or devices e.g., hospitals, financial or legal institutions. Typically, these organisations cannot risk stalling urgent patient care or legal proceedings and are forced to make difficult, time constrained decisions that have often resulted in the targeted organisations paying hefty ransoms to the attackers (O'Donnell, 2019). As more complicated versions of ransomware emerge, the more popular it becomes amongst cybercriminals. Ryuk is an example of a new and innovative strain

[37] Leakware and doxware are examples of ransomware that infect systems and obtain confidential information. If incriminating or damaging information is found, this arms the attacker with leverage to blackmail the victim. The attacker can subsequently threaten to publicize the damaging information unless a certain fee is paid.

[38] Leakware and doxware are not as commonly executed as crypto and locker ransomware. One reason for this is that the former strains of ransomware involve finding and extracting information. This process can be a more technically complex and more time-consuming process when compared to phishing spam that involves encryption (Fruhlinger, 2018).

of ransomware used to target low-tolerance organisations that encrypts data *and* internal back-up systems (as opposed to just data). Encrypting an organisations internal back-ups can force the hand of an organisation when the organisation was relying on reverting to internal back-ups in the event of a cybersecurity attack. A situation like this can cause the victim to believe that in order to get operations back up and running, they must pay the ransom.

While the generation of revenue via threat or extortion has been a long-established criminal practice, ransomware malware did not become popular cybercrime until after 2005 (Choi, et al., 2016). According to Ferreira & Kawakami (2018), ransomware is a new criminal offense extremely different from the extortion and kidnapping that society is used to. It poses as a new threat and it is necessary to create legislation that offers specific solutions to ransomware. In 2013, CryptoLocker ransomware emerged and due to its ability to rapidly spread infection in computers and networks in both private and public sectors, within a few months CryptoLocker garnered cybercriminals 27 million USD in ransom payments (Hampton & Baig, 2015). In the first quarter of 2016 there was an increase in ransomware victimisation by 3500% when compared to the fourth quarter of 2015 (Choi, et al., 2016). Fast forward three years and ransomware attacks were named by the Ponemon Institute as the fastest growing cybersecurity threat to organisations (Ponemon Institute, 2018). Security agency, Europol earmarked ransomware as one of the key malware threats in law enforcement (Europol, 2018). One key discovery from this report is the emergence and growth of ransomware-as-a-service i.e. the buying and selling of ransomware. Ransomware-as-a-service is an easy way for criminals to execute harmful attacks on vulnerable individuals or organisations (Europol, 2018). The perpetrators do not have to be technically adept to launch these types of attacks. They simply have to purchase the malware from a seller and launch their attack. This means that ransomware-as-a-service broadens the attack horizon for all criminals as no technical competence is required to execute an attack.

5.2.2 Impact & cost

Between 2017-2019, security company Symantec reported a multiplication in the number of targeted ransomware attacks, naming ransomware a “proliferating menace” and “a dangerous cybercrime threat facing organisations” (Symantec, 2019). The report points to marked growth in the number of criminal groups using ransomware, noting that some ransomware groups[39] are creating and executing ransomware themselves while others are selling ransomware as-a-service (Symantec, 2019). If we consider the period 2018-2019, Accenture recorded a 21% increase in the cost of attacks and a 15% increase in the number of attacks (Accenture, 2019). While another report from 2018 estimates that criminals garnered in excess of 1 billion USD from all ransomware extortions that involved the

[39] Ransomware groups have been affiliated to SamSam group, Ryuk, GoGaLocker (or LockerGoga), MegaCortex, RobbinHood, GandCrab and Crysis (Symantec, 2019).

encryption of data (McGuire, 2018). This figure is in fact expected to be much higher due to the increase of ransomware being sold to wannabe cyber criminals on the dark web which is reported as a different source of cybercrime, usually crimeware or Crime as a Service (CaaS).[40] While these reports suggest that the threat and cost of ransomware attacks is increasing, the exact amount that criminals are earning from ransomware attacks remains unknown. This is due to the fact that there are many ransomware groups who act together and others who act alone. Some attacks go under the radar as they go unreported by organisations for reasons including fear of the legal ramifications, the public cost of being party to an attack, a desire to maintain face, or protect intellectual property and corporate privacy (Marble, et al., 2015).

It is difficult to determine exactly how much a particular attack will cost the victim. The cost depends on a number of variables such as the type of ransomware executed, the organisation targeted, whether adequate external back-up systems are in place, how quickly the malware is detected, how quickly the malware spread, how many devices or systems have been infected. That being said, some data is available from security companies who are privy to organisations who have been attacked and the financial toll it took on them. For example, Coveware released their figures for the second quarter for 2019 and report that the average cost of a ransomware attack (which includes the ransom demanded, the cost of downtime and recovery (estimated to be 9.6 days at that time)) was 36,395 USD (Coveware, 2019b). This figure is a 184% increase on the first quarter of 2019 (Coveware, 2019b). In the last quarter of 2019, the average ransom demand increased to 84,116 USD with the maximum reported ransom being 780,000 USD (Coveware, 2019c). Fast forward to 2021, and the average cost of a ransomware payment in the first quarter of 2021 increased to USD 220,298 with twenty three days of expected downtime (Coveware, 2021). The total cost of a ransomware attack (including device and network cost, lost opportunity and ransom paid) in 2021 is averaging at US 1.85million (Sophos, 2021).

Specific strains of ransomware malware have much higher ransoms attached to them than others. For example, the average cost of Sodinokibi[41] (a popular choice of as ransomware as a service) is approximately 56,000 USD, Ryuk is 270,000 USD and Dharma is 14,000 USD (Coveware, 2019b). These three types of malware are used in targeted attacks and are much more sophisticated than the earlier 2005 Cryptolocker variants. LockerGoga is another successful and apparently costly ransomware that emerged in 2019. Yet, it stands out from the others for two main reasons. The first is that the attackers

[40] Ransomware is often bought or hired on crimeware platforms. The overall estimate for crimeware/CaaS (crime as a service) for 2018 was 1.6 billion USD (*See McGuire, 2018; 15).*

[41] The costs of Sodinokibi (also known as REvil) ransoms vary greatly. For example, while Coveware (2019b) reports the average cost as 56,000 USD, it is claimed that foreign exchange company Travelex paid 2.3million USD for the return of their sensitive files (including dates of birth, social security numbers of credit card data) that were obtained, deleted and ecrypted by Sodinokibi ransomware in 2020 (Isaac, et al., 2020; Spadafora, 2020).

do not state from the outset the amount they demand. Their ransom note states "The final price depends on how fast you contact us" (Rivero Lopez, 2019). The second is that LockerGoGa actually makes it difficult for the victim to pay the ransom as it logs the victim out of their system. As previously mentioned, black hat ransomware attacks are typically financially motivated. Yet, LockerGoGa is designed to stand in the way of the victim paying the ransom. This raises the possibility that ransomware attacks on enterprises are not always financially motivated. It reminds us that different attackers can have different goals and intentions. Furthermore, ransomware attacks can be used for other purposes than financial gain. For example, to make a political statement, socio-political gain, sabotage, terrorism etc.

5.2.3 Who is a target?

Few ransomware families like Ryuk are specifically designed to target larger enterprises, while the majority spread automatically and indiscriminately across the internet. In other words, ransomware attacks can be targeted or opportunistic and no industry has shown complete immunity from an attack. Considering black hats financial motivations, it is not surprising that targeted attacks on the financial services sectors continue to grow at exponential rates (Accenture, 2019). However, cybercriminals are also targeting other industries. For example, the NotPetya[42] ransomware attack targeted Danish transport giant Maersk in 2017 (Greenberg, 2018).[43] This attack had a devastating effect on businesses across Europe requiring a complete software infrastructure overhaul involving the reinstallation of thousands of machines (Osbourne, 2018). A successful attack in 2019 targeted Norwegian aluminium producer, Norsk Hydro. The attack shut down 22,000 computers across 170 sites in 40 countries and allegedly affected approximately 35,000 people (Tidy, 2019). Educational organisations as well as several municipalities in the United States were attacked in 2019 including New Bedford and Baltimore (Kaspersky, 2019b).[44] Even the healthcare sector has fallen victim to the exploits of ransomware attackers. For example, 2017's WannaCry exploit targeted a vulnerability within the National Health Service (NHS) in the United Kingdom (UK). This particular attack ably demonstrated the collateral damage that can ensue once ransomware has been initiated as the infection quickly spread across more than 150 countries crippling approximately 300,000 Windows computers (CyberArk, 2018). This

[42] NotPetya presents itself as ransomware however it could be considered wiper malware as it destroys data and disk structures. It appears that NotPetya attackers never intended to make the encrypted data recoverable (Department of Homeland Security, 2017; The Australian Cyber Security Centre (ACSC), et al., 2018)

[43] There is one key point about the ransomware that attacked Maersk. While NotPeyta ransomware resembled the ransomware Petya, the ransom message of NotPetya was only a ruse. The main goal of the malware was destruction i.e., to irreversibly encrypt a computer's master boot records. This is essentially the deep-seated part of the machine that tells it where to find its own operating system. No decryption key even existed. For more details see https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/

[44] Ryuk ransomware targeted New Bedford Massachusetts in July 2019 encrypting over 150 workstations and affecting 4% of the city's PCs (Muncaster, 2019). Ransomware strain, Robinhood, targeted Baltimore city in May 2019 (Krebs, 2019).

attack is believed to have affected 603 primary care facilities and other NHS organisations including 595 GP practices (National Audit Office, 2017).

5.2.4 How are victims responding?

One of the main problems for organisations who fall victim to an attack is that there is no legal or ethical guidance that states how an organisation should respond when they have been infected by ransomware. As a result, some victims are choosing to pay, others negotiate a lower fee, while a small few are choosing to not pay at all (Greenberg, 2018). It is possible to make a number of assumptions about why an organisation might choose to pay, negotiate, or not pay. For example, an organisation might choose to pay based on the idea that regaining access to critical files or devices as quickly as possible is the right decision for the business. Those who choose to negotiate might do so as they cannot afford to pay the original sum demanded but are open to paying a reduced fee. If it is determined that the data or devices captured are replaceable, organisations might decide to not pay as operations can continue as normal without any substantial interruption to services. Alternatively, an organisation might still decide to not pay based on moral grounds even if critical files and devices are encrypted or blocked. This was the case for Norsk Hydro, a business attacked by the LockerGoga ransomware in 2019 (Cimpanu, 2019) which we will discuss later in this text.

In cases where organisations choose to not pay the ransom, the cost of downtime appears to be much higher than those who choose to pay (Coveware, 2019b; Sjouwerman, 2018). For example, the Maersk and Norsk Hydro attacks were two targeted attacks where the victims chose to not pay the ransom demanded. It has been estimated that downtime respectively cost 200 million USD (Mathews, 2017) and 40 million USD (Ferguson, 2019). If the cost of the ransom is less than the cost of downtime, it is difficult to not jump to the conclusion that paying the ransom is the best decision for the organisation. On a superficial level, paying can have the facade of being the less expensive solution for organisations. Yet, this is not always the case.

By paying, there is no guarantee that a working decryption key will be provided – one may not exist or the received may be faulty. The former was the case in the Maersk attack as it came to light that a working decryption key did not exist. Thus, Maersk's decision to not pay a ransom paid-forward as paying the ransom would not have provided access to data or devices. This highlights one risk that comes with paying a ransom; the risk that a working decryption key does not exist. A second risk with paying is the risk of receiving a decryption key that is badly designed. Badly designed software can damage files during the decryption process which can arguably leave the organisation in a worse position than they were when the attack launched because they are out of pocket from paying the ransom and the damaged files are lost forever. This touches on a poignant ethical issue with paying or negotiating with criminals – how can an organisation trust that the ransomware attacker will do the right thing and supply a working decryption key when the ransom is paid? They cannot.

Another problem with paying or negotiating is that the organisation assumes the attacker's intention is monetary gain. An attacker can achieve continuous monetary gain if they develop a reputation as criminals who supply working decryption keys when ransoms are paid. While the irony of the situation is baffling, providing a reliable trustworthy service to victims of ransomware is one way of ensuring that victims will pay the ransom. In the alternative, where attackers develop a bad reputation as unreliable and untrustworthy criminals, organisations would presumably never choose to pay ransoms. The crux of the problem lies in the fact that not all ransomware attackers work off the same business model which means there is no guarantee access will be returned if a ransom is paid. A Coveware study estimates that 96% of victim organisations receive a working decryption tool after paying the ransom, while 4% do not (Coveware, 2019a). A more recent study conducted in 2021 specifies that 96% of those whose data was encrypted got their data back but only 65% of the those had their data restored (Sophos, 2021).

5.3 Ethical analysis

The case in consideration is a typical ransomware attack. The organisation has been attacked by a hacker and there is a loss of data and access to devices. A ransom is demanded. Which response to the attack could be considered ethically appropriate is ethically analysed. The unknowns include the attacker's intentions, how long downtime will last, and whether a decryption key exists.

Given that a ransomware attack is a time-sensitive situation where normative stakeholders short- and long-term interests can greatly vary, normative stakeholders' interests are separated into short- and long-term interests (where short-term refers to 0-6 months and long-term is 6-18 months post-ransomware attack). It is important to note that Phillips does not differentiate between short- and long-term interests. As ST is grounded in the concept of creating value for all stakeholders and that stakeholders' interests should be at the centre of managerial decision-making, the author is of the view that value can be created for stakeholders by considering stakeholders short- and long-term interests in cases where the managerial decision is likely to affect stakeholders short- or long-term interests differently. In such cases, short- and long-term interests ought to be considered in isolation. This critique of Phillips' method is intended to provide management with a holistic view of the impact the decision may have on legitimate stakeholders interests. Thus better positioning management to decide which action supports the continuation of the cooperative scheme. The determination of legitimate stakeholders is based on the voluntary acceptance to a mutually beneficial scheme of cooperation. It is possible that stakeholders may decide to remove their voluntary cooperation if management make a decision that conflicts with their short- or long-term interests. Equally, ethically minded stakeholders may be attracted to an organization and decide to engage in a mutually beneficial scheme of cooperation if an organisation makes decisions that satisfy their short-term interests.

5.3.1 Normative stakeholder interests

Normative stakeholders in this analysis include shareholders, employees, customers, suppliers, and the local community based on the assumption that these stakeholders are in a mutually beneficial cooperative scheme with the victim organisation that involves contribution and sacrifice and stand to be affected by the decision to pay, not pay or negotiate the ransom demanded. As Phillips explains, the organisation has an obligation of fairness to consider the interests of these stakeholders and as such, these stakeholders interests in the following sub-sections are considered (Phillips, 2003b: 217).

Shareholders' short term interests

Once the organisation has been attacked, shareholders have a financial interest in keeping the cost of the attack to a minimum. As previously mentioned, the average ransomware payment is 220,298 USD (Coveware, 2021) and the total cost of a ransomware attack (including device and network cost, lost opportunity and ransom paid) is averaging at 1.85million USD (Sophos, 2021). This is a significant amount of capital to lose unexpectedly. Paying the ransom might appear to be the cleaner-cut option when compared to the option to not pay. For example, not paying could potentially extend downtime beyond Coveware's estimate of twenty three days and incur a larger financial cost for the organisation. If we take this to be true, paying could be argued as being the cheaper, easier and less disruptive option. Choosing to pay could thus be argued as satisfying the short term interests of shareholders based on their interest in keeping costs at a minimum. However, paying the ransom does not guarantee that access to data or devices will be returned or that operations will immediately resume as normal. Only 65% of the encrypted data was restored after the ransom was paid (Sophos, 2021). If we compare the paying option to negotiating, paying outright may appear as the cheaper option in the short term as negotiating could increase downtime and costs for the organisation which would not be in the short term interests of shareholders.

Shareholders' long terms interests

Shareholders long-term economic interests in the organisation are motivated by their desire to get a return on their investment. This can only be regained from a growing, profitable and sustainable organisation. When an organisation chooses to pay a ransom, sustainability is at risk. This is due to the simple fact that by paying the ransom the organisation is actively contributing to the instability of cyberspace by financing criminal activity in cyberspace. Paying the ransom also proves to the attacker that it was a good business decision to attack the targeted organisation. In addition, the organisation affirms to the attackers that their business model is working, which in turn, is likely to encourage them to launch more attacks. The ransom money that the attacker receives from the organisation can be invested in new software to launch more sophisticated ransomware attacks. It could also be used to fund other nefarious activities such as organised crime or terrorism. From a Public Relations (PR) perspective, it is not in the long term interests of shareholders to be seen to be contributing to cybercrime or terrorism as this could damage the company's reputation. If we consider a situation where management outlined to their shareholders that their reputation is on the line if they decide to pay the ransom, it is reasonable to assume that paying the ransom may be unlikely to reach the communicative assent of the shareholders.

According to Editor in Chief of IT security researchers' website, *We Live Security*, organisations should not pay for the following reasons:

> "If you pay, you will support cybercrime activities by funding them with money; you don't have any guarantee that your information is going to be decrypted again. Remember, this is not a service, they are cybercriminals. [And] even if you pay, you are not going to be 'whitelisted' so you could get infected again so it's not a real solution for the future either. Prevention is the most important tool against Ransomware, since the infection can be usually cleaned afterwards but not always the information restored" (Thomas, 2015).

If it is true that paying a ransom increases an organisation's chances of being successfully attacked in the future and may result in an increase in the cost of cybersecurity implementation, repair and recovery services, paying the ransom could thus be argued as being in conflict with shareholders' long-term economic interests. If paying the ransom is in conflict with shareholders' long-term economic interests, paying the ransom does not support the continuation of the cooperative scheme. As the negotiate option can result in a similar outcome to the pay option, paying or negotiating are not in the long term interests of shareholders.

Employees' short terms interests

It is important for downtime to be contained and delimited for employees in the short term. This is mainly due to the immediate impact that the encryption of data or devices will have on working conditions for employees. For example, encryption can impede the fulfilment of daily tasks such as general administration and potentially high level activities depending on the organisation and the employee's role in the organisation. It can be assumed that it is in employees' short term interests for the attack to produce as little disruption for them as possible. If a working decryption key exists and the key is supplied upon payment, paying the ransom seems to remedy the impediment caused to employees' working conditions within the shortest timeframe possible. Despite negotiating having the potential to afford the organisation the chance of paying a lower ransom, it can delay the recovery process and the resumption of normal working conditions. Similarly, not paying the ransom at all can set back an organisation into weeks or months of recovery disrupting employees access to and the delivery of products or services.

Employees' long term interests

Generally, employees are interested in the long term success of the organisation, which can lead to employee salary increases, perks and bonuses as well as organisational growth and expansion. In addition, it can lead to securing permanent roles and offering increased opportunities for employees to advance to the upper echelons of the organisation. If the targeted organisation heavily relies upon cyberspace to execute their main functions, paying the ransom can threaten the sustainability of the very asset upon which their future depends. If paying the ransom undermines the security of cyberspace, brings a higher risk of being attacked, and results in increased cybersecurity costs in the future, paying the ransom does not appear to further the long term interests of employees - unless they are working for a firm that benefits from an unstable cyberenvironment and increased cybersecurity costs, e.g. a cybersecurity firm. There is some irony to this situation. Organisations who pay might do so as it appears to be the cheaper short term solution. However, by paying they could be potentially signing up for more costly cybersecurity protection against future attacks that they in fact have contributed to. The cost of paying a ransom coupled with the cost incurred from twenty three days of downtime as well as the potentially higher cost of cybersecurity protection in the future could directly impact employees' interests as it may result in reduced financial flexibility for the organisation to expand, offer employee promotions, salary increases or bonuses. Paying may also conflict with employees moral code and could result in employees confronting their employers. Beyond the direct impact this may have on the staff member and their relationship with their employer, this in turn could impact the organisation through loss of productivity or increased staff turnover. The same potential problems may arise in relation to negotiating. Consequently, paying or negotiating may not be in the long term interests of employees.

Customers' short term interests

Depending on the type of organisation targeted and the products or services offered, it is likely that a short disruption to services versus a long-term disruption might be a determining factor for customers. It is reasonable to assume that the shorter the downtime time, the more likely it is that customers will wait for access to return. The longer the downtime, the higher the risk that a customer will suffer from a loss of data access to the organization if the organisation was hosting information or a service on behalf of the customer, e.g. a cloud storage or service provider. If a working decryption key exists and is supplied once the ransom is paid, paying the ransom could be argued as being in the short term interests of the customer.

Customers' long term interests

A study by Wenger et al. (2017) shows that EU citizens particularly value the privacy and safety of data in cyberspace. As paying contributes to cybercriminal activity which can include anything from fraud, identity theft as well as ransomware attacks, it does not appear to be in the long-term interests of European customers who are concerned about the privacy and safety of their data. Subsequently, paying is in conflict with customers' long term interests.

The organisation might have cybersecurity insurance which will require the organisation, if they are covered, to pay an excess which will influence the organisation's premium for the following year. As previously mentioned, ransom money sent to the attackers can encourage more attacks. In the long term, this can create a vicious cycle of paying out ransoms and increasing the frequency of attacks which can cause the cost of cybersecurity insurance to increase. Higher costs for organisations may force organisations to rebalance their accounts to compensate for growing cybersecurity expenses. This could involve increasing the price of products or services, which is not in the long term interests of customers interested in getting value for money or purchasing products and services at competitive prices.

Suppliers' short term interests

The suppliers who might be affected by a ransomware attack can include numerous commercial entities. They can range from outsourced consultants such as IT specialists to outsourced services such as CSP. Depending on the type and scale of the attack, supplier's devices and data could be infected. As suppliers have economic interests in the organisation, a speedy return to full operation is paramount in the short term. As previously alluded to, paying does not guarantee a return of access but it does appear to satisfy a supplier's short-term interest in resuming operations as quickly as possible.

Suppliers' long term interests

Suppliers have a long-term interest in continuing their relationship with the organisation with whom they are in a mutually beneficial scheme of cooperation. Due to the interconnected nature of cyberspace, paying the ransom can be harmful to suppliers in the long term for the same reasons that it can be harmful to shareholders, customers and employees – it threatens the security of cyberspace. It is rational for all functioning organisations including suppliers to not want to dilute the security of cyberspace, contribute to cybercrime and increase the cost of cybersecurity services or insurance. It is in suppliers' long-term interests to improve the security of cyberspace as it is the tool which they heavily rely upon to execute their daily functions. This suggests that not paying is in the long-term interests of suppliers.

Local communities' short term interests

Local communities try to build relationships with organisations with the aim of attracting capital from investors to set up in their area. The goal of local communities is to increase services and the quality of life of local residents. It is therefore of paramount importance to local communities that the organisations already located in their community minimise the cost and collateral damage caused by any cybersecurity attack in the short term. The pay option could thus be viewed as being in the short-term interests of the local community.

Local communities' long-term interests

Certain local communities have economic interests that stand to be affected if they invested in the affected organisation by offering them lower tax rates to set up in their region. In such cases, the local community will wish to minimise the cost and damage the attack incurs in order to regain their investment through local employment[45] and the development of services. If the organisation chooses to respond to a ransomware attack by paying, this may affect the reputation of the local community and other organisations within the area. For example, if it becomes commonplace in one area for organisations to pay attackers, this is not a good reflection on the cybersecurity practices adopted by businesses in the area. It could also be a reflection of the quality of the cybersecurity practices implemented by employees from the local community, e.g., insufficiently attentive to good practice or being in cahoots with attackers. This may deter future investors and is not something that a local community would want. Generally speaking, local communities have a long-term interest in creating an attractive reputation at home and abroad to attract future investors to their area. They also have a long-term interest in the sustainability of the firms currently within their locality. It could hence be argued that an organisation paying the ransom is not in the long-term interests of the local community.

5.3.2 Is there a difference between paying and negotiating?

When compared to paying outright, negotiating could be viewed as an attempt to pay a lesser amount towards cybercrime. Interestingly, attempts to negotiate with ransomware attackers might work. For example, an investigation conducted by F-Secure found that organisations can bargain an average discount of 27% on the original sum demanded if they enter negotiations with the attackers (Michael,

[45] Generally, employees reside within close proximity to their workplace thus the interests of employees and the local community are closely connected.

2016). The same study found that ransomware groups have developed customer friendly websites that enable victims to easily access, navigate through and make swift payments. Some even go as far as having a section for "Frequently Asked Questions". These customer focused features are something akin to what a legitimate business would call a "customer journey" – starting when the victim realizes they have been attacked, all the way up to converting the victim into a paying customer (Michael, 2016). Providing such a "reliable" and "trustworthy" service can be misleading and may convince victims that they can trust that the attacker will provide a working decryption key when the ransom has been paid.

There appear to be two main problems with negotiating. The first problem is that negotiating catalyses the same vicious feedback loop as paying outright – money is exchanged with cybercriminals which creates the same conflict of interests between normative stakeholders short and long term interests. The second problem - also present in the pay scenario - is that negotiating can be a costly affair as it prolongs downtime. Yet, the option of negotiating may be favourable to the targeted organisation. For example, the organisation could use the time during this period to find the decryption key themselves if they have the competency and means to do so. If the organisation cannot find a working decryption key and are not in a position to not pay the ransom, the organisation could then proceed with the negotiation process and pay a lower fee to the attackers. Negotiating can take a longer time than the option of immediately paying and there are benefits to this. Negotiating gives the organisation more time to thoroughly examine the responses available to them and discuss their planned recovery strategy. This additional time could be used to determine what the ransomware is doing or whether a decryption key is already available. If it is unlikely that a working key exists or will be in any way useful, the organisation could make an informed decision about whether to proceed with the negotiation process or decide to not pay at all. As in the Maersk attack, after an investigation, the organisation discovered that the ransomware was deleting and corrupting files making them irrecoverable with or without a decryption key. Negotiating can also allow for the introduction of law enforcement agencies. Many agencies have cybersecurity teams who are highly competent cybersecurity experts who might be able to offer some assistance in finding the key, deciphering the ransomware used etc. The difference between entering the negotiation process and choosing to not pay is that the immediate decision to not pay allows the organisation to immediately take action and initiate its recovery process straight away. A delayed decision might impede recovery and prolong downtime and may result in higher costs for the organisation in the long run, unless negotiating is used as an intentional stalling tactic.

5.3.3 Derivative stakeholder interests

Phillips states that it is in the interests of the organisation for management to spend a limited amount of time and resources attending to derivative stakeholders (Phillips, 2003b: 222). In order to stay true to Phillips' approach, the interests of derivative stakeholders are deliberated upon. Derivative stakeholders are named as competitors, the news media, the attacker and law enforcement.

Competitors' interests

Competitors are actors who are in the same market sector as the targeted organisation. Naturally, competitors have a commercial interest in the cybersecurity threat landscape specific to their industry. Being privy to such information enables competitors to formulate a strategy against the most prevalent threats to their industry. Competitors have a financial interest in knowing the average cost of cybersecurity attacks in their industry. In addition, they have in interest in knowing the average time it can take to recover from an attack so they can integrate this data into their recovery strategy. Competitors will also be interested in knowing the 'industry standard': is it to pay, not pay or negotiate ransoms? And how likely is it for the organisation to get their information back if they do decide to pay the ransom?

There is always the possibility that all organisations within one industry band together and decide to never pay ransomware money. If this happened over a long period of time, it would become clear to ransomware attackers that hacking that particular industry has a low pay off when compared to other industries who continue to pay ransoms and the attackers incentive to attack that industry would reduce. The cumulative impact of a whole industry deciding to not pay could therefore be profound, not only in terms of reducing the frequency of ransomware attacks but also in terms of reducing cybersecurity costs and cybersecurity insurance rates for that industry. On the other hand, if there is an industry-wide precedent to pay ransoms or negotiate ransoms, this could cause any intuitive ransomware attacker to target that specific industry as they are almost guaranteed a pay-out. It is hence in the long term interests of competitors for organisations to not pay.

News Media interests

Currently, the media is saturated with reports of cybersecurity attacks on organisations. It has become commonplace for readers to see that yet another organisation has been hit by a ransomware attack and the organisation has chosen to pay or negotiate a lower fee to get their data back. However, there are few documented articles about organisations who choose to not pay ransoms, why they chose to not pay and what are the financial and operational consequences of not paying. If the sole interest of the media is to grab the attention of their readers, it is reasonable to assume that reporting atypical behaviour is in their interests. Therefore, an 'out of the ordinary' organisational response to a ransomware attack such as not paying the ransom, could interest the media. In the same breath, it is equally fair to assume that the media have an interest in reporting the most controversial story which could involve an organisation paying an extortionate ransom and never receiving a working decryption key or negotiating a lower price and getting a sizeable reduction on the initial ransom demanded.

Regardless of whether the organisation chooses to pay, not pay or negotiate, it is in the media's interest to cover a story regardless of whether it is favourable or damaging to an organisation. In addition, the organisation has an obligation of fairness to their normative stakeholders to appease the news media in situations where their coverage of events could be extremely damaging to the organisation and its normative stakeholders. Depending on the relationship with the media and the affected organisation, it may be in the media's interests to publish or not publish the story.

Attacker's interests

Managing threat from adversaries can involve putting procedures, software and hardware in place to mitigate the harm that attackers can inflict on the organisation. Let us assume that the attacker's sole interest in this situation is monetary gain. In this case, the attacker wants the highest ransom fee to be paid by the targeted business. The higher the ransom paid, the more likely it is that black hats can continue their line of 'business'. The attacker's interests will be thus satisfied if the organisation chooses to pay the initial amount demanded or negotiate a lower fee. Both scenarios encourage the attacker to continue their efforts of extortion. The decision to not pay the attacker, however, switches the paradigm. The organisation does not relinquish control to the attacker by paying them for the return of what is rightly theirs, nor does management contribute to weakening cyber space or more specifically, cyber security. Organisations who choose to pay or at the very least negotiate a lower fee to ransomware attackers satisfies the interests of the attacker.

Law enforcement's interests

We name law enforcement as the final derivative stakeholder in this situation. The organisation has an obligation of fairness to their normative stakeholders to abide by the law. For example, to not commit fraud, to file accurate tax returns and to satisfy other legal obligations, as failing to do so is illegal and can negatively impact the interests of the organisation and its normative stakeholders. While the organisation does not owe the legal system any additional moral obligation of fairness to advance their interests when they are choosing how to respond to a ransomware attack, they have an obligation of fairness to advance the interests of their normative stakeholders by taking the time to consider the interests of law enforcement. For example, one goal of law enforcement is to improve the security of cyberspace. Paying ransoms directly undermines law enforcement's attempt to achieve this goal as paying directly contributes to cybercrime and other nefarious activities. This has the potential to create a more volatile cyber environment for the organisation and its normative stakeholders – both parties to whom law enforcement have an interest in protecting. Given that negotiating involves paying a lower sum to cybercriminals, it is a payment all the same which is not in law enforcement's interests. The only situation that appears to satisfy law enforcements interests is the not pay scenario.[46]

5.3.4 Prioritising conflicting stakeholder interests

The analysis of stakeholder's interests suggests that conflict does not arise between the interests of different normative stakeholders. Although, it does arise between their interests in the short and long term. This is due to all normative stakeholders having similar short- and long-terms interests. On the one hand, normative stakeholders collectively have a short term interest in resuming business as quickly as possible. On the other, they have a long term interest in contributing to a safer cyberenvironment as a volatile cyber environment can negatively affect the continuation of the cooperative scheme.

Should normative stakeholders' short term interests be prioritized over their long-term interests, or the other way around?[47] Phillips suggests that conflict amongst stakeholders can be adjudicated by choosing an action that supports the continuation of the cooperative scheme and is likely to reach the communicative assent of all normative stakeholders. While this is not a conflict amongst stakeholders but one between terms, it still leaves management to decide which action is ethically appropriate. Phillips does not cover such an instance arising in his research but in this circumstance, it would be remiss to overlook the conflict between short and long terms interests. To determine which of the two interests should be prioritised, Phillips' conceptual approach to managing conflict amongst stakeholders is applied.

If management choose to pay the ransom, this offers some potential advantages. If a working decryption key exists and the attacker supplies a decryption key upon receipt of payment, there will be minimal downtime (estimated to be twenty three days) which will be followed by the assumed resumption of working conditions for employees and suppliers and access to services for customers. The disadvantages of paying include not regaining access to the encrypted data because it has been deleted, corrupted or a decryption key never existed to begin with. In such a scenario, downtime can be indefinite. For the purposes of illustration, let us now imagine how management might frame the pay option to their normative stakeholders. It could be something like this:

[46] One might also argue that payment for criminal behaviour, albeit as a victim, could never be regarded as legal under the general legal principle of benefiting from crime.

[47] Mari (2009) argues that long-term benefits trump short-term benefits in sustainble work systems.

Paying is in all normative stakeholders' short terms interests as it is likely to allow us to resume business as quickly as possible. However, it requires us to trust that the cybercriminals are honest and will uphold their end of the deal. Paying will involve reallocating finances to pay the ransom. While there is a small risk of renege on the part of the attacker, there is a high probability that access will be returned. Paying however, does not exempt us from future attacks. Paying, in fact, increases our chances of being attacked at some point in the future as the ransom money is likely to be used to fund further attacks and weaken the security of cyberspace. Paying also reaffirms to the hackers that there is a likely chance, if we are attacked again, that we will pay in future which may be a factor in planning future attacks. If other organisations follow our example, it is likely that paying will contribute to increasing the costs of cybersecurity insurance within our industry which may impact services and our ability to expand or offer promotions, benefits to employees.

The negotiate option may be very similarly framed by management. The one main difference being that management pay a lower, negotiated ransom. Negotiation can bring about the similar advantages and disadvantage as the pay option based on the assumption that a decryption key exists and it will be supplied when a fee is paid except for one minor difference. The difference is that the organisation contributes a lesser amount to cybercrime by paying a reduced fee which enables management to present itself as being proactive in attempting to minimize the damage. Management can thus demonstrate to shareholders and other stakeholders (to whom this is revealed) a higher level of management competence in their attention to stakeholder needs than if they had just paid immediately. A second difference between negotiating and paying is that business resumption may be delayed due to the negotiation period.

The practical advantages of not paying are obviously polar to those listed for the paying and negotiate scenarios. Management may frame this option to the normative stakeholders in the following way:
Making a quick decision to not pay allows us to immediately implement our recovery plan (assuming that management have an incident response plan in place). We can reallocate funds to rebuild workstations, servers and begin installing necessary software to re-establish operations. There will be a loss of business and there will be a time/cost associated with recollecting and/or regenerating the data. The integrity of the organisation is unlikely to be tainted as we chose to not contribute to cybercrime or exchange any money with cyber criminals. We can communicate the details of the attack and our choice to not pay with the authorities, selected suppliers/customers, competitors, cybersecurity specialists which can contribute to industry-wide defences. We may also communicate this information to the news media as our experience may influence others to respond in the same way should they find themselves in a similar position.

It would be remiss to overlook the possibility that there can be exceptional cases in which choosing to pay might be in the long-term interests of the normative stakeholders, might reach the communicative assent of all stakeholders and might support the continuation of the cooperative scheme. One example where this situation can arise is when the targeted organisation is in a financially precarious position prior to the attack. If not paying the ransom may expediate the dissolution of the organisation, not paying could be viewed as being an action that is not in the long-term interests of normative stakeholders. An action that catalyses the dissolution of the organisation, does not support the continuation of the cooperative scheme. An action that does not support the continuation of the cooperative scheme is unlikely to reach the communicative assent of all stakeholders.

5.4 Conclusion

This conceptual stakeholder analysis suggests that not paying a ransom will satisfy the long-term interests of all normative stakeholders and complement the continuation of the cooperative scheme. Not paying the ransom is also likely to reach the communicative assent of all stakeholders. Given that one requirement of conflict resolution in Phillips's approach is that the proposed action must support the continuation of the cooperative scheme, it is fair to assume that long term interests trump short-term interests. On that basis, the ethically appropriate response to a ransomware attack is to not pay the ransom. However, in rare circumstances paying a ransom or negotiating a lower fee will be in the long-terms interests of all normative stakeholders. This is likely to occur when not paying the ransom may expedite the dissolution of the firm or risk causing serious harm to the organisation and its stakeholders. Therefore, the conclusion of this analysis is that the decision to pay, not pay or negotiate is case dependent and every case should be critically analysed by management to determine the ethically appropriate response.

Chapter 6 Attempting a Botnet Takedown

6.1 Introduction

This chapter analyses the ethics of a botnet takedown in response a DDoS attack. A DDoS attack is a "malicious attempt using multiple systems to make a computer or network resources unavailable to its intended users, usually by interrupting or suspending services connected to the Internet" (Revuelto, et al., 2017: 1). DDoS attacks are a growing threat to organisations and involve overwhelming a system with a large volume of traffic which can interrupt internet-based services (Kaspersky, 2020). DDoS attacks are primarily used by adversaries to target organisations by disrupting internet connections, slow the performance of a website or take it offline (Kelly, et al., 2020; Russell, 2017; Sucuri, 2019). Attempting to disable the network from which a DDoS attack is staged is known as a botnet takedown. Botnets are insecure devices connected to the internet that are remotely controlled by adversaries enabling them to orchestrate large DDoS attacks. A botnet takedown is one of many active cybersecurity responses organisations can adopt in response to the ever-growing threat of DDoS attacks (Himma, 2004). As far back as 2004, security company Symbiot recognised a need to adopt more aggressive tactics in response to DDoS attacks and launched a tool that had the capability of profiling and blacklisting Internet Service Providers (ISPs) as well as launching counter DDoS attacks (Kotadia, 2004).

Active responses – also known as Active Cyber Defence (ACD) – to DDoS attacks vary in terms of their impact on the attackers network. For example, some can cause minor, temporary or excessive disruption while others can permanently damage the attacker's network (Pattinson, 2020). International organisations such as the European Cybersecurity Emergency Response Team (CERT-EU) advocate ACD responses like DNS[48] sinkholing (Revuelto, et al., 2017). DNS sinkholing works by intercepting a DNS request that is attempting to connect to a known malicious domain[49] and returning a controlled IP address (ENISA, 2021b). The controlled IP address directs the malicious domain to a sinkhole server. This technique helps prevent hosts from connecting to or communicating with known malicious destinations such as a botnet Command and Control (C&C)[50] server. DNS sinkholing enables researchers to isolate and analyse the redirected malicious web traffic without causing much disruption to the attackers' network (Kaimal, et al., 2019). It is inexpensive to set up and maintain,

[48] DNS is a service that allows internet users to access a website by name (called domain name) rather than by Internet Protocol (IP) address. IP addresses are difficult to remember as they consist of four numbers containing one to three digits which are separated by dots i.e., 69.63.176.13. This means that domain names simply act as a link to the IP address allowing the user to enter the name of a website e.g., Facebook into a search engine and access the website without needing to remember the IP address.

[49] There are lists of known malicious domains that sinkhole administrators can use to create a sinkhole.

[50] Bots report to C&C servers and it is through C&C servers that criminals can control bots and give them orders. In the absence of C&C servers, bots are useless (ENISA, 2021a).

however this approach does not cure the infected systems, nor does it deter the attacker from launching future attacks.

Unlike DNS sinkholing, a botnet takedown can be used to disable the network from which the attack is staged (Pattinson, 2020). Although less conventional than DNS sinkholing, successful botnet takedowns involve the removal of the C&C server rendering the entire botnet useless. While a successful botnet takedown does not cure the infected systems, it shuts down the botnet and presents the possibility of deterring future attacks (Hoffman & Levite, 2017). Shutting down a botnet is possible with the cooperation of an ISP when there is only one server and the location is known. ISP involvement is advantageous as ISPs can restrict access to certain resources by disconnecting or limiting access to the internet (Aasmann, 2011). When ISPs and victim organisations work together in a botnet takedown, the organisations and individuals whose devices were used as bots to launch a DDoS attack can be contacted and instructed as to the steps they can take to appropriately clean their infected device. As an innocent person's devices are manipulated by the botmaster to both amplify the power of the DDoS attack and obfuscate the attacker's location and identity, it is very likely that during the botnet takedown, the functionality of these devices is further impacted (Hoffman & Levite, 2017; Pattinson, 2020). This temporary disruption may cause harm to the device owner. As a consequence of this potential harm caused to innocent third parties, botnet takedowns are considered a controversial and aggressive (yet not the most aggressive[51]) form of ACD (Pattinson, 2020).

Botnet takedowns are of particular interest for three reasons: 1) DDoS attacks are a growing threat to organisations 2) a lack of prosecution has resulted in organisations taking matters into their own hands and 3) the topic of attempting a botnet takedown is lacking from business ethics literature.

1) DDoS attacks are a growing threat to organisations and there are no indicators that suggest the impact, cost and frequency of DDoS attacks will decrease in the near future (see section 6.2.1 for more details). Impact, cost and frequency are influenced by the ever-growing mass distribution of insecure devices wired to or wirelessly connected to the internet. When infected by malicious software, these devices can target many systems and networks (see section 6.2.2 for more details).

2) A lack of prosecution in the area of cybercrime has resulted in organisations taking matters into their own hands. Some organisations are exploring the option to counterattack a DDoS attack by

[51] Hacking back is typically considered the most aggressive form of ACD, one that is typically undertaken with the intention of causing permanent damage or destruction to the attackers network (Hoffman & Levite, 2017; Pattinson, 2020).

attempting a botnet takedown to track, trace and respond to the perpetrators themselves, both with and without the assistance of law enforcement (see section 6.2.3 for more details).

3) Passive responses to DDoS attacks offer very little protection once under attack. As an alternative, or in addition to existing defences, organisations can choose from a range of active responses. Ethical literature tends to refer to the wide-ranging category of active responses as ACDs and broadly categorizes ACD as either wholly ethical or unethical despite the variances and complexities associated with each active response (see section 6.2.3 for more details). An over-simplification like this can misrepresent the complexity of each response. The aim of this chapter is to solely focus on one active response – botnet takedowns – and determine the ethicality of this response by analysing the impact it may have on legitimate stakeholders (see section 6.3 for more details). For example, the decision to counterstrike a DDoS attack with an attempted botnet takedown presents the victim organisation with the potential to prevent the attacker from launching more attacks from the same source and may deter the attacker from launching future attacks. However, there are risks when attempting a botnet takedown, such as disrupting systems owned by innocent individuals. It is also possible that the attacker may launch subsequent, more aggressive attacks in retaliation to the attempted takedown. It is thus worth considering in more depth the risks associated with an attempted botnet takedown as well as how those risks may impact the interests of the main stakeholders in the victim organisation (see section 6.3 for more details).

This chapter is structured in the following way. It begins with a description of the relevance and growing threat that DDoS attacks present to organisations including their impact and cost (Section 6.2.1). The relationship between DoS, DDoS and botnets is discussed and a brief explanation of the different types of DDoS attacks is provided (Section 6.2.2). As this is an ethical analysis of attempting botnet takedowns, existing active defences described by the literature are discussed (Section 6.2.3). Phillips' stakeholder approach described in Chapter 3 is used to systematically analyse stakeholders' interests and how they may be affected by an attempted botnet takedown (Section 6.3), leading to a conclusion in Section 6.4.

6.2 DDoS attacks

6.2.1 A growing threat to organisations

The first DDoS event occurred in 1988 when a self-replicating computer program written by Robert Morris (known as the Morris worm) collected host, network and user information (Spafford, 1989). Whilst the proliferating worm was designed to identify network users, it spread quickly, disrupting normal activities and internet connectivity on 10% of the 60,000 machines connected to the precursor internet ARPANET (The Advanced Research Projects Agency Network) (FBI, 2018). The first malicious DDoS attack occurred in 1996 targeting Panix, an ISP, affecting the operations of commercial institutions (Keromytis, 2011). The attack caused large losses to victims and was executed by flooding electronic networks with traffic they could not handle, knocking them offline (BCS, 2017). In 1997 the University of Minnesota was targeted, causing the university's network to go offline (Radware, 2017). As a result, service was denied to legitimate users including students, suppliers and staff members for more than two days.

The early 2000s witnessed several media outlets and large MNCs being targeted including CNN, Dell, E-Trade, eBay, Amazon and Yahoo! (BCS, 2017). In 2007, government services and financial institutions in Estonia fell victim to DDoS attacks which were considered the first acts of cyber warfare, as it came in response to a political conflict with Russia over the relocation of the 'Bronze Soldier of Talinn', a World War II Statue (CloudFlare, 2020). By 2012, social networking platforms like Facebook, Twitter and YouTube and other institutions like the Bank of America, Citibank, HSBC and London's Internet exchange all suffered from a large-scale DDoS attack (BCS, 2017; CloudFlare, 2020).

In 2016, the largest on-record DDoS attack occurred against DNS provider, Dyn (CloudFlare, 2020). This attack is particularly notable as it rendered a significant portion of the internet inoperable, leaving many high-profile web services unreachable for several hours (Mansfield-Devine, 2016). The Mirai malware was the primary source of the malicious attack as it created a botnet out of compromised IoT devices such as cameras, smart TVs, radios, printers and baby monitors (Kelly, et al., 2020; Ko, et al., 2020; Kambourakis, et al., 2017). The attack involved 100,000 compromised devices and cultivated a strength of 1.2 terabits per second (Tbps) (CloudFlare, 2020). As a consequence, the operations of many large MNCS were disrupted including Deutsche Telekom, taking 900,000 of their customers offline (Russell, 2017). Other affected organisations included Airbnb, Netflix, PayPal, Visa, Amazon, The New York Times, Reddit and GitHub.

Criminal group DDoS for Bitcoin ("DD4BC") emerged in 2014 targeting the online gambling industry and other entertainment, financial and energy-based companies (Singapore CERT, 2020). DD4BC are known for sending an email notification to the target informing them that they are under attack by way of a low-level DDoS attack. A ransom is then demanded in Bitcoin and in return DD4BC will abstain from launching a larger DDoS attack (Bisson, 2015). This is not the first instance of adversaries using DDoS attacks as an instrument of extortion. As mentioned in Section 1.3, DDoS attacks are often executed in conjunction with other attack vectors such as ransomware. Ransom DDoS (RDDoS) attacks have existed since the late 1990s (Singapore CERT, 2020). They usually begin when the attacker threatens to launch a DDoS attack unless a ransom is paid within a given timeframe. Either before or after the threatening message is delivered, the attacker will often demonstrate the credibility and validity of their threat by launching demonstrative DDoS attacks.

In October 2020, an incident involving a RDDoS attack was executed via the SunCrypt ransomware. Once the target organisation became aware that they were under attack, they engaged in negotiations with the attackers. When negotiations stalled, the adversaries launched a DDoS attack. When the victim organisation logged into the Tor payment site, the display message stated that should the victim wish for the DDoS attack to stop, they must continue negotiations and pay the agreed ransom (Abrams, 2020). Thousands of global organisations across various sectors have been targeted by RDDoS attacks, one of whom was the New Zealand Stock Exchange (NZX) who suffered trading disruption for several days as their ISP, Spark, was repeatedly targeted by RDDoS attacks (Singapore CERT, 2020).

Moreover, DDoS attacks can be used as smokescreens for more nefarious network infiltrations (Henderson, 2017). For example, they can be used to divert the attention of the victim organisation by executing low-bandwidth, sub-saturating DDoS attacks. Such attacks are extremely common and, as they usually last a short period of time (less than ten minutes), they do not cause extended periods of downtime and can appear harmless (Newman, 2019). The objective of these small DDoS attacks (less than 10 Gbps) is not to cripple the target organisation's website, but rather to be disruptive enough to take firewalls and Intrusion Prevention Systems (IPS) offline, leaving the network open to attack. During this short time, adversaries can use automated scanning or penetration techniques to target, map and infiltrate a network to steal data or install malware such as ransomware. These short DDoS attacks often go undetected by defence systems (Newman, 2019). This allows adversaries to launch various short, targeted attacks to test for vulnerabilities within a network and monitor the success of new methods without being detected. This means that adversaries can experiment with new techniques before deploying them at scale.

Launching DDoS attacks is not limited to highly competent attackers. Any individual or organisation with a financial or ideological motive for example, can cheaply wield the power of DDoS attacks. This is due to the widespread availability of ready-made, prepacked DDoS packages that are sold as DDoS-as-a-service. For example, DDoS-as-a-service can be bought for 10 USD per hour to target an unprotected website, launching 10-50,000 requests per second (Mission Critical, 2020). Alternatively, a potential buyer can pay 60 USD for a 24-hour attack (Gomez, 2020). If it turns out that the targeted website has premium protection, it is possible to purchase a DDoS attack that launches 20-50,000 requests per second and uses multiple elite proxies for 200 USD for 24 hours (Gomez, 2020).

Given the variety of attack vectors used in conjunction with DDoS attacks as well as the widespread availability of DDoS-as-a-service, it is not surprising that the number of DDoS attacks continue to soar in frequency and size (IDG, 2018).

Security experts Kaspersky believe that the increase in the frequency of attacks during the year 2020 is partly due to the COVID-19 pandemic (Kaspersky, 2020) as the pandemic forced organisations to embrace remote working (Kaspersky, 2020; NetScout, 2020). Organisations were unprepared to support a fully remote workforce, weakening their security posture due to changes in work and infrastructure patterns (ENISA, 2020b). Organisational security risks increase as online home usage increases the risk of malware being installed, especially when employees are using the same network as roommates, spouses, or children. Connecting to insecure networks or using insecure devices e.g. mobile phones also increases organisational security risks. By taking advantage of security vulnerabilities, cybercriminals successfully targeted internet-dependent organisations such as ecommerce business, educational platforms and financial services (NetScout, 2020). For the month of May 2020 alone, NetScout reported 929,000 DDoS attacks (NetScout, 2020), while security company Lumen reported a monthly increase of 1200% in emergency DDoS mitigation activations between the months of July and October in 2020 (Lumen, 2020). Two ENISA[52] reports published in 2020 also highlight the staggering increase in the number of DDoS attacks on organisations (ENISA, 2020a; ENISA, 2020b). One of which compared the third quarter of 2019 with the same quarter in 2018 and found a 241% increase in the total number of DDoS attacks (ENISA, 2020a). A more recent study conducted by Nexusguard reported a 278% increase in DDoS attacks in the second quarter of 2020 when compared to the same period in 2019. This was a 542% increase compared to the previous quarter (Nexusguard, 2020).

In respect of attackers' goals, it appears that malicious actors are motivated by their intention to cause as much disruption as possible (Imperva, 2019). This is evidenced in Imperva's 2019 report where two thirds of targets were persistently attacked up to five times and a quarter of targets were attacked ten times or more (Imperva, 2019). The most-attacked industries according to the number of attacks and number of targets were the gaming and gambling sectors, closely followed by computer- and internet-based organisations and, thereafter, the financial sector (Imperva, 2019). In addition to their relentless persistence, malicious actors are advancing their technical skills, using more robust defence methods, using new vectors such as reflection (sometimes referred to as amplification[53]) attacks to target organisations and advertising DDoS-as-a-service on common social media channels like YouTube and

[52] ENISA is the European Union Agency for Cybersecurity, an agency dedicated to achieving a common level of cybersecurity across Europe. The agency contributes towards cyber policy, awareness and training.

[53] Amplification attacks are reflection-based volumetric DDoS attacks which can be executed when an attacker overwhelms a target server or network with an amplified amount of traffic. As a result, the targeted server and the surrounding infrastructure become inaccessible (Cloudfare, 2021).

Reddit (as opposed to the historical practice of advertising DDoS-as-a-service on the dark web) (ENISA, 2020a). These are contributing factors that have led to a progressive rise in the power and popularity of DDoS attacks. For example, the largest recorded DDoS attack at the time of writing was in 2020. The attack attained a power of 2.3 terabits per second (Tbps) dwarfing previous record-breaking attacks against GitHub in 2018 and Dyn DNS server in 2016 (McCallion, 2020).

The potential interruption caused by DDoS attacks is not only inconvenient for organisations, it can also disrupt operations and functions for hours or weeks (Jayaswal, et al., 2002). For example, the longest recorded attack duration in 2019 was 138 days (TSA, 2020). Moreover, DDoS attacks can be expensive to remediate (Kaimal, et al., 2019). The exact cost varies depending on the number of DDoS attacks executed, the characteristics of the business targeted and the environment in which the attack occurs (IDG, 2018). Lerner (2014) estimates that a single minute of network downtime (the time it takes to get services back online) can cost an organisation on average 5,660 USD (which accumulates to 300,000 USD per hour). While these numbers suggest that the longer the attack duration, the higher the cost for organisations, this number can greatly vary depending on the organisation involved.

6.2.2 DoS, DDoS, botnets & takedowns

DoS attacks are different from DDoS attacks as they stem from one source and a single internet connection – they are not geographically distributed. DoS attacks are usually used to target one computer system and tend to be much smaller than DDoS attacks with a typical DoS attack generating tens of Megabits per second (Mbps) of malicious traffic (Grimes, 2017). DoS attacks are more akin to one-on-one combat where the attacker must have a higher bandwidth than the victim in order to be successful. This limits the potential targets of the attacker as the attacker would not be capable of successfully launching a DoS attack on a large enterprise with higher bandwidth (Grimes, 2017). This restriction is circumvented in DDoS attacks because the adversary can simply create fake traffic to amplify the power of the attack to execute a large-scale distributed attack. DDoS attacks tend to be much larger than DoS attacks and typically start in the hundreds of megabits per second. They can target multiple internet connections at a time, stem from multiple sources and locations, and can damage several systems at once. The ability to generate more power has made DDoS attacks a more popular and powerful version of DoS attacks (Jayaswal, et al., 2002).

Adversaries can create fake traffic by taking advantage of insecure IoT. The IoT is a large network of millions of devices that are capable of interacting with one another and transferring data without human intervention (Vishwakarma & Jain, 2019). Such devices include smart cameras, light switches, voice controllers, doorbell cameras, wireless sensors, computers, mobile phones, baby monitors, medical and industrial equipment (Gartner, 2015). It was estimated that the 9.9 billion active IoT devices in 2020 will increase worldwide to 21.5 billion by 2025 (Statista, 2018). The problem that IoT devices create for organisations is that a large majority of them are not properly secured by manufacturers before they go to market (Kelly, et al., 2020). This makes them vulnerable to existing malware lurking on the internet which has led to the IoT devices being coined the Internet of Vulnerabilities (Angrishi, 2017).

Compromising an insecure device can be achieved by remotely infecting it with malware bots. A malware bot is a software application or script that performs automated tasks on command. An adversary controls a bot at will (known as the master controller) by sending commands to the infected device. The device becomes a 'zombie computer' performing tasks received from the master controller which often go undetected by the device owner/user (Williams, 2015). Adversaries herd huge numbers of zombie computers and network them so they can all be controlled at once to perform a large-scale DDoS attack. This network of bots is called a botnet (Zetter, 2015). Bots are effectively an army of computerised robots on standby waiting for instructions from the master controller (Kelly, et al., 2020). It is the proliferation of the IoT devices and the generation of botnets that are the primary factors that enable DDoS attacks to be so successful (Angrishi, 2017).

There are four basic methods attackers can use to deploy the C&C server. They include direct, centralised, decentralised using Peer-to-Peer (P2P) communication, and hybrid. A direct C&C mechanism is the easiest to set up. It allows the bot-master to directly control, recruit and disseminate commands to the botnets and individual bots from the C&C server. This architecture is the least resilient to a takedown out of the four as it can be easy to trace commands back to the point of origin. However, it is growing in popularity due to the availability of sophisticated machine identity obfuscation techniques. The most common type is centralised which communicates via Internet Relay Chat (IRC). IRC networks can continually switch channels to avoid being taken down and are widely used to host botnets, coordinate DDoS attacks and spam campaigns (Kashinath, 2021).

Decentralised architecture is different from both direct and centralised architectures as it is self-sufficient. P2P bot-code is engineered in such a way that infected devices can serve as a bot as well as a C&C server to at least one other connected device. In other words, the bot code can recruit new bots to join the botnet and inject the new bots with the C&C capability. As a consequence, there is no single point of failure, making P2P mechanisms quite resilient to take down measures. The final architecture, hybrid, is the most difficult to set up out of the four architectures due to the sophistication of the bot code. It is an amalgamation of the strengths of direct, P2P and centralised botnets. It offers better resilience and is more difficult to take out. This is based on the very small chance that the machine which the commands are traced back to is actually the originating attackers' machine. For context, obscuring the location of the attacker in a DDoS attack is like forging a return address on a letter in order to conceal the sender's identity (Sucuri, 2019). This is achieved when the bot master sends IP packets with a fake source IP address. This makes DDoS attacks more difficult to identify, trace and block when compared to direct DoS attacks which are easy to detect, attribute and mitigate (Grimes, 2017). The most common type, centralised, is considered in the example analysis in Section 6.3.

The life cycle of a botnet involves various phases. Yimu & Shangdong (2019) describe six. The first involves the botnet spreading through different viruses or worms. The second phase is where the bot downloads the entire botnet program. The third entails the bot contacting the C&C server. Once the bot is authenticated, the bot joins the botnet group and the fourth phase is complete. In the second to last phase, the bot starts receiving commands from the C&C server and in the final stage, the C&C server will launch or stop an attack (Yimu & Shangdong, 2019).

It is beyond the remit of this analysis to explain every type of DDoS attack. For the sake of simplicity and context, the most commonly executed attack type is SYN Floods otherwise known as SYN attacks. SYN attacks accounted for 80% of all DDoS attacks in 2020 (ENISA, 2020b) and they were named in the top three[54] global DDoS attack types in 2019 (TSA, 2020). SYN attacks work by exploiting the way in which a connection[55] is established between a user and a server. In a SYN attack, the attacker repeatedly sends initial connection request (SYN) packets to the target device. The target device responds to each request by leaving a port open and ready to establish a connection. However, the connection is never established as the final packet called an ACK packet never arrives. Meanwhile, the attacker continues to send more SYN packets to the target until all the ports are in-use and unavailable. As a result, the server becomes overwhelmed and is not able to function properly, causing the server to respond to legitimate traffic slowly or not at all (CloudFlare, 2020b).

6.2.3 Defence methods: From passive to active

Governments across Europe are attempting to better position themselves in the fight against cybercrime. For example, the Dutch government proposed new legislation that would give police agencies the power to hack into computers and install spyware (Sarhan, et al., 2018). European counterparts such as Germany, Poland, France and the UK have been using hacking tools such as FinSpy to secretly monitor and surveil criminals' keystrokes (a tool that does this is called a keylogger) and

[54] The other two attack vectors listed were total traffic attacks and User Datagram Protocol (UDP) attacks (TSA, 2020).

[55] Establishing a connection between a user and server involves three steps. These steps are known as the "handshake" process of a Transmission Control Protocol (TCP) connection. In normal circumstances, the first step of establishing a connection requires a user to send a SYN packet to the server to initiate a connection. The second step involves the server responding with a SYN/ACK packet to acknowledge the communication. The final step requires the user to return an ACK packet in acknowledgement of receipt of the packet from the server. The result of the three steps is an open TCP connection through which the bot can send and receive data.

access audio, camera and screenshot tools (European Parliament, 2017). Despite having these hacking tools to hand, the likelihood of law enforcement prosecuting cybercriminals is less than 0.05% (World Economic Forum, 2020). In fact, cybercrime is a flourishing to such an extent that it is estimated that by 2021 the cybercrime industry will be equivalent to the Gross Domestic Product (GDP) of Japan, the third largest economy in the world, which is worth 6 trillion USD (World Economic Forum, 2020).

Ajayi (2016) explains that there are a number of reasons why cybercrime remains an albatross. They include the issue of tracing adversaries who use tools like Tor to obfuscate their identity. If a criminal is identified but situated outside the jurisdiction where the victim domiciles, there is a lack of international law that compels sovereign nations to return cybercriminals for trial. In combination with the lack of adequate legislation, the lack of effective legislation where extant, and the existence of international law without enforcement mechanisms are contributing factors that impede a crack down on cybercrime. The cost of investigating cybercrime on top of poorly trained, low-paid law enforcement agencies that lack protection are also contributing factors (Ajayi, 2016).

Due to a lack of prosecution, law enforcement agencies have altered their approach from punishment to denial of service. They are collaborating with private entities to seize and disable a botnet rather than prosecute the master controller. For example, in 2011 the FBI and the United States Justice Department worked together to hijack and eliminate the Coreflood Botnet. The US government initiated and won a civil suit in federal court seeking a temporary restraining order allowing it to replace servers, collect IP addresses and deliver a disabling command (Zetter, 2011). A second example of law enforcement agencies using denial in the absence of arrest or prosecution is when various organisations including the FBI, US Marshalls and the digital crimes unit from Microsoft worked together to identify and disrupt the Citadel botnet (Lerner, 2014). Citadel installed key logging software on infected computers, which enabled the master controller to track everything the user typed. The denial operation involved sinkholing in which servers were set up to mimic the botnet C&C and collect IP addresses of the zombies. Once identified, the owners of the zombie computers and the ISPs were contacted and offered step-by-step instructions on how to remove the botnet. As a result, communications were stopped between 1462 separate botnets operated by Citadel and the millions of devices infected by them (Lerner, 2014). Even though the effort did not result in the arrest and prosecution of the alleged Citadel master, it did result in the disruption of over 90% of the Citadel botnet (Lerner, 2014). A third example is the cooperative effort between Facebook and the FBI taking down the Lecpetex botnet which was being used by cybercriminals to steal Facebook and other credentials from 250,000 infected machines (Mimossa, 2014). A more recent example is the Microsoft's collaborative attempt with the U.S. military's Cyber Command to disable Trickbot (Burt, 2020).

Organisations are also trying to grapple with the threat of DDoS attacks. There are currently no passive defence measures organisations can adopt that offer complete protection from DDoS attacks (Crane, et al., 2013; Kaimal, et al., 2019; Watkins, et al., 2015). As alluded to in Section 4.2.2, passive measures include observing traffic on the network, using firewalls as well as monitoring for surges in traffic and checking for packet flooding. Organisations can outsource to a cloud-based DDoS prevention solution e.g. Incapsula or Cisco Solutions or other DDoS mitigation providers like Arbor, AKAMAI, A10, Radware and Neustar (Revuelto, et al., 2017). They can employ a red team to simulate a DDoS attack to establish how prepared the organisation is should they come under attack, identify problem areas and develop a mitigation plan. They can use a Content Delivery Network (CDN) which is a whole network of proxy servers to filter traffic coming into the website, block harmful DDoS traffic and protect the origin server's IP while allowing legitimate users to access the website (Jaiswal, 2020). Alternatively, organisations can pay for more bandwidth than they will ever really need (called bandwidth overprovision). In the event of a DDoS attack, this buys the blue team time it to detect irregular traffic and take action before the website shuts down.

Many commercial victims worry about the effects that publicizing an attack might have on their relationships with customers and fear that customers may become alarmed after learning of security breaches and may respond by taking their business elsewhere (Dittrich & Himma, 2006). Firms may choose to minimize such deleterious effects by responding internally to digital intrusions without involving the media or law enforcement. Many private actors lack access to the manpower required to internally manage digital intrusions as well as the sophisticated attribution tools and information available to the government however some private actors possess these resources (Holzer & Lerums, 2016). Actively responding in the form of Active Cyber Defence (ACD), as a substitute for involving law enforcement agencies may be reasoned as a viable option when the victim organisation believes that it is far more efficient to manage the issue themselves. This is particularly the case when victim organisations believe that law enforcement do not have adequate resources to respond to digital intrusions, cannot keep pace with the frequency or severity of attacks, and/or have very low arrest and prosecution rates (Dittrich & Himma, 2006).

Organisations, security professionals, academics and private individuals are already exploring and practising active responses to attacks. For example, a 2012 survey of Information Security Professionals indicated that 36% of the 181 participants engaged in retaliatory hacking 23% said they had done this on one occasion and 13% admitted to doing it frequently (US-CERT, 2012). Crane, et al. (2013) identified a counter-attack cyber defence tool called booby trapping that enables victims to directly react to attacks. They claim that the tool offers a direct and automatic response to a detected intrusion by automatically tracking and locking down malicious traffic in addition to identifying the ISPs hosting the attacker's controlled servers. Joshi & Goudar (2014) proposed a spyware program that has the capability of bypassing any antivirus software or firewall on the hackers device which enables the victim to get the user's details and attack the hacker (Joshi & Goudar, 2014: 239). Watkins, et al., (2015) describe an active defence method that can be used to stop in-progress DDoS attacks that utilise the Dirt Jumper Family (DJF) of botnets.

Academics[56] discuss ACD in a range of ways including whether ACD is legal, whether ACD is an ethical cyberwarfare tactic or whether ACD is ethical under ethical principles such as the Necessity Principle. For example, Himma (2004) considers the issue of whether it is ethically permissible for private persons or entities to adopt aggressive ACD measures in response to general hacker attacks under the Necessity Principle. The Necessity Principle allows one person to infringe the rights of an innocent person if doing so is necessary to achieve a significantly greater moral good. Himma (2004) acknowledges that we cannot reliably predict the effects, direct or indirect, of the most aggressive measures on innocent parties and concludes that it is hard to determine whether the moral good involved in infringing such rights outweighs the moral costs involved. He states, "if a victim of a hacker attack wishes to adopt these measures, she will have to justify them under some other commonly accepted ethical principle" (Himma, 2004:39).

Dittrich & Himma (2005) describe five levels of ACD responses to cybersecurity attacks based on wireless information warfare. The response levels range from passive responses that involve minimal engagement on the part of the victim to counter-striking which entails reaching beyond the resources owned and operated by the victim in an effort to identify, mitigate or eliminate the threat. An example of responding to a DDoS attack with proportional force is provided and organisations are discouraged from doing so as it may adversely affect the organisation's own interests should the counter-strike result in escalation. Dittrich & Himma (2005) conclude that the growing frequency of hacker attacks

[56] A search of the Scopus database using search string "cybersecurity", "hack", "back" yielded three results on 11 October 2020. Two relevant papers and one irrelevant. The relevant papers (Mihelič & Vrhovec (2018) and Shackelford, et al., (2019) are included in the body of this chapter. A search string, 'active' 'defence' and 'cybersecurity' in the scopus database yielded the same two papers. A third search on the Scopus database using search string 'cyber', 'security', 'hack', 'back' yielded seven results: Two irrelevant and a book chapter that the author was unable to access. From a review of the abstract, it appears that the chapter focuses on approaches used by governments to defend national interests. The remaining four results are included in this chapter (Bradbury, 2013; Kallberg, 2015; Shackelford, et al., 2019; Watkins, et al., 2015;). Snowballing widened the relevant literature. As did a further search of the Web of Science database using search string "cyber', 'security", "hack", "back". This yielded fourteen results. Five are relevant and included in this chapter (Denning, 2013; Holzer & Lerums, 2016; Pattinson, 2020; Shackelford, et al., 2019; Watkins, et al., 2015). These sources were also snowballed for further relevant publications.

combined with the increasing inability of law enforcement agencies to respond adequately has created a need for a coordinated active response that involves public and private institutions.

Denning (2008) analyses ACD from the viewpoint of armed conflict including cyber warfare at three levels: state level, hacktivism conducted by non-state actors and active response. She creates a framework that analyses whether an attack resembles force and whether an attack follows the principles of the law of war. In a more recent publication Denning (2013) reviews the concepts of active and passive air and missile defences and applies them to cyberspace. Messerschmidt (2013) argues that international law should govern cross-border hacks, in particular hack-backs by private actors as a proportionate counter measure to an attack. Kallberg (2015) lists attribution (where the victim organisation may not be capable of identifying the attacker) and the fact that deterring future attacks is not guaranteed as problems with hacking back. He also raises the possibility that the attacker may launch a second strike which can result in uncontrolled escalation. The issue of attribution is also noted by Holzer & Lerums (2016) who analyse the legal complexities of hacking back across international waters from two perspectives. The first is hacking back from a criminal perspective where hacking back is a criminal act. The second is from a military perspective where hacking back is viewed as an acceptable military style response to an act of aggression.[57]

Hoffman & Levite (2017) attest that private organisations are increasingly complementing their passive cybersecurity pratices with ACD to meet the demand for effective defence that is currently not being provided by governments. They advocate for a principle-based approach encouraging the practice of private sector ACD as a complement to passive measures. They claim that although an active approach brings risk, if executed by bounding principles and industry models, it has the potential for long-term, cumulative benefits. Rather than narrowly focusing on the risks or pitfalls of one single technique or capability, there should be a proactive attempt to minimise the cumulative risks of ACD and look at the impact of ACD from a holistic perspective because doing so could substantially improve private sector defence and change the calculus of malicious actors (Hoffman & Levite, 2017).

Denning & Strawser (2017) compare ACD using analogies from defence, arguing that ACD is neither offensive nor necessarily harmful or dangerous. It can be executed in accordance with well-established ethical principles relating to harm, necessity and proportionality. Specifically, in relation to botnet takedowns, they state "active defences mitigate substantial harm" Denning & Strawser (2017: 74).

[57] The appropriate response depends on the nature of the initial act and its potential results (Holzer & Lerums, 2016).

Mihelič & Vrhovec's (2018) research focuses on employing hacking back to protect and secure critical infrastructure.

According to Neal (2019), cybervictims are resorting to revenge and retaliation hackbacks against cyberattackers due to the lack of effective cybersecurity guidance and deterrence by law enforcers. Neals concludes that there is a need for ACD as a tactical and strategic option for deterrence. While Shackelford, et al., (2019) notes firms like FireEye publicly admitting to hacking back and suggests that any policy that permits ACD "should be narrowly tailored to only allow passive active defense measures under strict government oversight, and only then for the worst cyber attacks on civilian critical infrastructure sectors" (Shackelford, et al., 2019: 427).

Pattinson (2020) argues that even though private cybersecurity firms engage in offensive operations to infiltrate, disrupt, and destroy the systems of actual or potential aggressors, they should not be permitted to perform offensive forms of ACD such as hacking back due to the myriad problems it creates. Other research echoes a similar determination in respect of hacking back (Stevens, 2020). Stevens analyses hacking back from the perspective of a 'right to self defence' and claims that counter strikes in cyberspace are not equivalent to self defence and should not be encouraged.

It is clear from the above that ACD is broadly referred to as wholly ethical or unethical in the ethical literature. There is a significant lack of focus on the ethicality of specific types of ACD such as attempting a botnet takedown. In addition, there is even less coverage of attempting a botnet takedown from a business ethics perspective. In the proceeding section, the focus is solely on attempting a botnet takedown from a business ethics perspective.

6.3 Ethical analysis

6.3.1 Normative stakeholders' interests

This section is a stakeholder analysis of one form of ACD, attempting a botnet takedown. The case in consideration is an organisation that has been successfully targeted by a centralised DDoS attack. The organisation has the time, technical capability and resources to cooperate with an ISP who can hijack bot nodes and contact the owners of the bots with instructions to remove the offending malware. Phillips' stakeholder analysis method as described in Chapter 3 is applied in order to determine the interests of normative and derivative stakeholders relative to the decision to be made which enables the conceptual determination of whether conflicts of interests arise. Based on normative stakeholder interests, it is possible to determine whether an attempted botnet takedown is or is not likely to achieve the communicative assent of all stakeholders. If the proposed action is likely to achieve the communicative assent of all normative stakeholders and supports the continuation of the cooperative scheme, the proposed action will be considered ethically appropriate.

Normative stakeholders according to Phillips are groups or individuals who are in a mutually beneficial cooperative scheme with the victim organisation that involves contribution and sacrifice. Normative stakeholders also include those who stand to be affected by the decision to be made. Phillips explains that the organisation has an obligation of fairness to consider the interests of normative stakeholders (2003b: 217). In this analysis, the victim organisation is the organisation that has suffered a DDoS attack and is faced with the decision of whether to attempt a botnet takedown. Groups and individuals who are in a beneficial cooperative scheme that involves contribution and sacrifice and who stand to be affected by the decision to attempt a botnet takedown include shareholders, employees, customers, suppliers and the local community. The interests of these normative stakeholders are considered in detail in the proceeding subsections.

Shareholders' interests

As explained in Section 5.3.1, shareholders have interests in stock prices and the value of the organisation they have invested in. DDoS attacks negatively impact stock prices causing them to drop. If services to customers are interrupted there can be a delayed negative impact on market value. While this suggests that the indirect economic impact of a DDoS attack is context-dependent, it is fair to assume that shareholders have a legitimate financial interest in avoiding any interruption of services to customers in order to limit the financial cost of a DDoS attack. Attempting a botnet takedown from which the attack is staged presents an opportunity to prevent the same attack from happening again. This is relevant in cases where an organisation is targeted more than once by the same botnet. If the successful disablement of a botnet can limit future attacks from the same source and thus reduce the likelihood of services to customers being disrupted, shareholders have an interest in the successful takedown of botnets. However, it is possible that an attempted botnet takedown does not guarantee success. There is no meta-analysis available that states ACD will definitively result in either success, failure or otherwise or what success or failure looks like. For example, is success the dismantling of a botnet or deterrence or both? This means that in the absence of data it is unlikely that shareholders will be able to make a quantified risk assessment. Hence why the possibility of deterrence and the possibility of non-deterrence are raised as unknowns in this analysis. Attempting a botnet takedown is a reactive measure which means that while it is possible that a successful botnet takedown may stop a similar assault, the successful first attack is unlikely to be stopped by a botnet takedown (because of the time it takes to track and disable the attack). Furthermore, the first attack can be effective enough to cause hours of downtime and disruption to the organisation.

As cybercriminal activity continues to grow, so too do cyber claims and cyber insurance premiums. European insurance claims managed by Marsh Continental Europe state that cyber claims grew by 83% between 2018 and 2019 with 67% of all claims being malicious and 28% being accidental (Consultancy.eu, 2020). This growing rate of cyber incidents and claims is attributed to an over-reliance on digital channels such as AI and cloud technology as well as the IoT which has expanded cyber risks for many organisations (Consultancy.eu, 2020). It is also attributed to cyber-attacks being a cheap and risk-free way for criminals to export money from companies. Data collected in a 2019 Cyber Claims Study (which compiled 2,081 cyber claims from American, Canadian and British organisations) indicates an increase in the percentage of cyber claims caused by criminal activity (Net Diligence, 2019). In a three-year period, the percentage increased from 72% in 2014 to 86% in 2017. The same study shows that the median claim for successfully targeted SMEs[58] who were unable to continue operations due to either ransomware or denial was 49,000 USD. For larger companies the median cost was 1.7 million USD. It is in shareholders' interests for all cybercriminal activity to reduce now more than ever because the success of modern organisations heavily depends upon digital infrastructure. In addition, adversaries are doubling-up on the mechanisms they use to disrupt business operations such as RDDoS attacks. A reduction in cyber-criminal activity will reduce the number of cyber claims caused by malicious actors. This can influence the cost of cyber insurance premiums which are affected by a number of factors including a company's industry, data risks, exposures, computer and network security and annual gross revenue (Marciano, 2020). Successful botnet takedowns are considered a tactical strategic option for deterrence (Neal, 2019). If an attempted takedown could deter future cybercriminal activity, the action would be in shareholders' financial interests. It has been argued that active defences like successful botnet takedowns have cumulative benefits, minimise risks and improve private sector defence (Hoffman & Levite, 2017). Yet, the same problem remains; success is not guaranteed. If success is not guaranteed, failure is a possibility. A failed takedown does not represent shareholders' financial interests in reducing malicious cyber claims and cyber insurance premiums.

Preventing a second attack and limiting disruption to customer services are the ideal outcomes that an attempted botnet takedown can produce. However, botnets are very versatile and adaptive, and adversaries are continuously employing state-of-the-art techniques to evade detection (Chang, et al., 2015). TrickBot is a prime example of a complex botnet that has infected over a million computing devices. Microsoft formed an international alliance with a number of organisations including ESET, Symantec and Lumen's Black Lotus Labs to stop the spread of Trickbot (Burt, 2020). The alliance identified the precise IP addresses of the C&C servers that the infected devices received instructions

[58] Organisations with an annual revenue of less than 2 billion USD (Net Diligence, 2019)

from, produced evidence of same in court and were granted approval from the Eastern District Court of Virginia to suspend all services to stop Trickbot operations (Burt, 2020). The suspension included disabling the IP addresses and rendering the content stored on the C&C servers inaccessible. On 12 October 2020 Microsoft announced that it managed to legally disable Trickbot. However, they did not dismantle it completely as one week after the announcement was made, a newer version of the Trojan surfaced (Arghire, 2020). This case highlights that an attempted takedown does not guarantee that the botnet will be successfully stopped. There is no guarantee that the attacker will be caught, brought to justice i.e., arrested and/or prosecuted, or deterred by the counterstrike. In fact, this case illustrates that it is possible that the attacker could be reinvigorated by the attempted takedown to adapt and update the botnet to evade further detection by private organisations and/or law enforcement. This case also illustrates that an attempted takedown can require a concerted collaboration with other private entities e.g., ISPs, website owners and law enforcers. The fact that the cybercrime industry is comparable in value to the Japanese economy suggests that a collaborative effort to take down botnets is a warranted endeavour. Yet, the cumulative power and resources generated from collaborations to date are not always successful (as seen from the Trickbot takedown mentioned above). The only apparent certainty of a failed takedown, either as one entity or as part of an alliance, is that it will require a significant amount of planning and organisation – an endeavour that can be time-consuming and resource intensive.

An alternative to updating the botnet to circumvent detection, the attacker may decide to launch a subsequent and more aggressive attack on the organisation in response to the attempted takedown. This may harm corporate image and cause further damage to the organisation and its stakeholders. Choosing a reactive measure like a botnet takedown in response to a DDoS that provokes the attacker into launching an additional, more harmful attack is not in shareholders' interests.

Attempting a botnet takedown may raise potential legal and political issues if the takedown involves disrupting external networks. Depending on the jurisdiction in which the attempt is undertaken, the organisation may find themselves subject to legal proceedings. An eventuality that is not in shareholders' interests.

It is a real possibility that shareholders in the affected organisation are institutional and have an interest in many different businesses. For those shareholders it will be in their interests for the organisation to adopt tactics that deter future attacks with the aim of improving the security of cyberspace and reducing cybercrime. Similarly, it will be in these shareholders' interests for each organisation in which they are invested to limit collateral damage when under attack and resume business as quickly as possible.

The above analysis suggests that a successful botnet takedown is in the interests of shareholders as it may stem the likelihood of an attack being staged from the same source. A successful takedown may also deter future attacks, reduce cybercrime and malicious cyber insurance claims and cyber insurance premiums. It is equally evident that an attempted takedown comes with various risks i.e., it does not guarantee success, is a resource intensive and time-consuming endeavour, could result in legal proceedings for the organisation and could provoke the attacker to either reinvent the existing botnet to evade future detection or launch more harmful attacks. There are no metrics that indicate the success or failure rate of attempted botnet takedowns. In addition, there are so many variables at play it is not reasonable to state conclusively whether adopting a ACD approach like attempting a botnet takedown will in every case be in shareholders' interests. It is clear that shareholders' interests hinge on the chance that the attempt will result in success as the consequences of a failed attempt include wasting valuable time and resources that do not result in the detection, prosecution or arrest of the attacker. This is in addition to the risk of escalation, provoking the attacker into reinventing the botnet and causing harm to innocent third parties who may subsequently decide to take legal action against the organisation. It is evident that if these risks eventuate, shareholders' financial interests in avoiding both escalation and causing harm to third parties will not be satisfied.

Employees' interests

Employees may double-up as shareholders. This means it is possible that employees will have similar interests to shareholders who are not employees. For example, limiting the impact the attack has on stock prices and the value of the organisation. General employees' interests also align with common shareholders' interests as listed above. For example, employees have an interest in limiting disruption to customer services and resuming business as quickly as possible. If a large number of customers migrate to a competitor, this may impact the organisation's ability to retain future business. If there is no customer, there is no business. Such a situation will negatively affect the organisation's financial flexibility to offer permanent employment. Depending on the severity of the interruption to customer services and the associated reputational damage, the organisation could be forced to make redundancies.

Limiting disruption also extends to employees' ability to complete daily tasks that may have been interrupted due to the attack. This is particularly relevant to employees who work in an industry that stores, processes, collects, analyses, shares and/or delivers data, products or services online. For example, e-marketing, advertising, e-Learning, e-commerce, social media platforms, online gaming or gambling. Even more basic functions like sending and receiving emails, managing and accessing inventory may be disrupted by a DDoS attack. If disruption impacts employee performance this could be raised at employee performance reviews. Employees thus have an interest in avoiding escalation and limiting collateral damage that may delay the resumption of business and impact the organisation’s financial flexibility to use finances for employee-centric activities such as offering benefits or perks, bonuses or promotions.

Employees are primarily concerned with maintaining employment and retaining income. They thus have an interest in working in an organisation that makes concerted efforts to practice activities promote prosperity and sustainability. It is likely that employees are not privy to management reasoning to engage or pass on actions or decisions. Therefore, employees may not hear about the organisation attempting a botnet takedown until it is made public, leaked or heard through the grapevine. Organisations who appear to not err of the side of caution and appear to make risky decisions that threaten the long-term success of the organisation may deter employees from working there. The same applies to organisations that engage in legally and ethically questionable behaviour. Employees have an interest in working for an organisation that can withstand, respond and bounce back from a DDoS attack. As mentioned in the shareholder section, a successful botnet takedown may prevent a second DDoS attack staged from the same source, but prevention is not guaranteed. An unsuccessful botnet takedown that consumes a significant amount of time and resources may be perceived as counterproductive to employees who are cognizant that those resources could have been invested elsewhere i.e., improving the security posture of the organisation.

Similar to shareholders interests, employees’ interest in attempting a botnet takedown hinges on whether the attempt is a success. From an employee’s perspective, an attempted takedown may be considered a success when the successful dismantlement of the botnet resulted in deterring future attacks and reducing the costs of cybercrime insurance and premiums. Success would also include avoiding escalation, the business could promptly return to normal working conditions, customer base was not greatly impacted, reputational damage was limited, and employees were able to fulfil tasks in such a way that employee performance was not significantly affected. A failed attempt is almost the opposite to a successful attempt. For example, an attempt contributes towards the volatility of cyberspace which can impact corporate image, decrease the value of the organisation and increase the price of cybercrime premiums. In cases where the attempted takedown invigorates the attacker to launch subsequent attacks, this is likely to increase the cost of the initial attack which may, in the worst-case scenario, affect the organisation’s ability to retain permanent employees. It is thus clear that the successful takedown of a botnet is in employees’ interests and a failed attempt is not.

Customers' interests

In the event of a DDoS attack, customers of the victim organisation can double up as owners of the infected devices used to amplify the DDoS attack (a possibility that is in fact, applicable to all normative stakeholders). This may be an intentional or unintentional act from the attacker. From a practical point of view, attackers might have a marginal interest in intentionally using customer machines. It is marginal as this would require the attacker to differentiate attack traffic from regular customer traffic. Whereas traffic from a non-customer or traffic that accidentally includes a customer's device is more likely to be utilised to target an organisation as it is less challenging. As previously explained, an infected device is considered compromised when malware has been installed and is being remotely controlled by the attacker (Keromytis, 2011). A compromised device such as a computer will show signs of infection in the guise of slower performance than usual, slow programs and applications, reduced internet speed, the display screen may frequently freeze, or the device may not shut down properly (Williams, 2015). Device owners i.e., customers will have an interest in manufacturers embedding adequate protection in IoT devices before they go to market. However, as some manufacturers are remiss in their efforts to do this, devices are easily manipulated by attackers to launch attacks like DDoS attacks.

Device owners can practice good cyber hygiene and update their device as advised to do so by the respective OS vendor. Yet, good cyber hygiene is not common amongst device owners (see Annex B) and good cyber hygiene does not necessarily protect a device owner from a DDoS attack as DDoS attacks exploit an innate vulnerability in the internet's infrastructure. This raises the question, if manufacturers are selling insecure devices, who is responsible for securing them? Is it the device owner, law enforcement or some other entity? Answering this question is beyond the remit of this chapter but it is a pertinent question to raise in this discussion as it highlights the fact that the buck is being passed and there is currently no organisation or institution who are offering a fast and effective solution. It is in device owner's interests to receive a notification from a relevant authority (whether it is from the organisation who suffered from a DDoS attack, ISP or OS vendor) that their device is being used as a bot and receive the appropriate advice on how to clean the infected device.

Botnets can transcend various jurisdictions. Given that there is no universal law on attempting takedowns, legislation differs from country to country, state to state. This means that attempting a botnet takedown without input from authorities from all affected regions, could result in legal proceedings. This points to the benefit of collective action. Collective action has been previously utilised by OS vendors like Microsoft and law enforcers to takedown botnets. Working collaboratively has enabled organisations such as ISPs, OS vendors and legal authorities to identify compromised devices within the realms of the law. In the absence of collective botnet takedowns (where the organisation internally tries to attempt a takedown on their own), tracing the infected devices is likely to infringe state law and is unlikely to be successful without consultation from an ISP or OS vendor. It is thus, in customers', shareholders' and suppliers' interests for organisations to collectively act in conjunction with law enforcers and suppliers, e.g. OS vendors and ISPs, particularly when customers are unable, or do not know how to protect their devices themselves.

Some organisations believe that it is far more efficient to respond to cybersecurity attacks internally without the involvement of law enforcement (Dittrich & Himma, 2006). One reason why an organisation may choose to circumvent authorities is to avoid the additional time it takes to get legislative approval as well as the reputational damage associated with the public announcement of the attack which is likely to occur once authorities have been notified. However, engaging in an illegal botnet takedown may result in losing or gaining customer trust. According to Wenger et al, (2017), losing the trust and confidence of customers is considered one of the most damaging consequences of cybersecurity attacks to businesses. If customer trust is lost due to illegal activity, customers may take their business to a competitor which will exacerbate the reputational and financial damage created from the initial attack. It is equally possible that customer trust is gained when customers view the active response as a necessary endeavour in the absence of prosecuting cybercriminals. Customers may see the internal response as appropriate as it protects customers interests and improves the security of cyberspace. This shift in perspective could sway customers from wanting to disengage after hearing of the initial attack to wanting to remain a customer. This perspective shift is likely to be influenced by the outcome of the attempted takedown as the advantages of a failed botnet takedown are less likely to be apparent to the average customer. While the disadvantages of a failed attempt will be more obvious to the average customer such as consuming time and resources. One advantage of a failed takedown is that it sends a message to attackers that organisations are fighting back.

If collective action is taken, it is in customers' interests for the organisation to be aware of the risks, particularly those that may cause further disruption to customer services. Botnet takedowns can involve extended periods of time and often involve law enforcement which can delay the takedown process (Dittrich, 2015). For example, Torpig (otherwise known as Mebroot, Sinowall and Anserin), Ozdok and Pushdo/Cutwail, Mariposa, Bredolab and Coreflood takedowns all involved law enforcement and/or an extensive civil legal process (Dittrich, 2015). Only a handful of botnets have been successfully reported as taken down or taken over by private organisations without the aid of law enforcers i.e., Waledac, Rustock, Kelihos and Zeus (Dittrich, 2015). If a botnet takedown causes a further delay to products or services, an attempted takedown is not in the interests of customers who want to experience as little disruption as possible. Attempting a botnet takedown may incur higher recovery costs forcing the organisation to re-balance their accounts to compensate for the expense incurred. An increase in cybersecurity related costs could affect the organisation’s ability to offer products and services at a competitive price. It is in customers' interests for the organisation to have the financial flexibility to offer competitive prices.

The above analysis suggests that an attempted botnet takedown does not satisfy customers’ interests in resuming business as quickly as possible. While a successful botnet takedown may satisfy customers’ interests in identifying an infected device and may satisfy customers’ interests in accessing products and services in a safe way (if the successful takedown results in deterrence and a reduction in cybercrime), success is not guaranteed. In fact, the attempt may delay access to products and services if escalation occurs which is in conflict with customers’ interests. Therefore, an attempted takedown that is successful and results in deterrence and a reduction in cybercrime is in customers’ interests. An unsuccessful takedown does not satisfy customers’ interests.

Suppliers’ interests

Access to suppliers’ services can be disrupted by a DDoS attack especially when those services are managed through an online platform. Suppliers will have an economic interest in the organisation speedily returning to full operation and limiting any disruption to the product(s) or service(s) the suppliers provide to the organisation. Suppliers who rely on cyberspace to conduct business will have an interest in the organisation taking measures that improve the security of cyberspace and deter the efforts of cybercriminals. A successful botnet takedown presents the opportunity to disrupt criminal operations and deter future attacks. This opportunity comes with the risk of escalation, a delayed recovery and the potential to cause harm to innocent third parties including customers and suppliers. However, a successful or unsuccessful attempted takedown that provokes a retaliatory attack is in conflict with suppliers’ economic interests in the organisation making a speedy recovery. A retaliatory attack is also in conflict with supplier’s interest in improving the security of cyberspace and keeping cybersecurity costs down.

Suppliers such as Microsoft Windows, Apple macOS, Linux, Android and Apple’s iOS are common OS vendors. As previously mentioned, Microsoft have played a leading role working with national authorities in efforts to track, trace and dismantle harmful botnets. OS vendors can aid the detection and removal of malware on devices that run on their software. Therefore, OS vendors can help at the back end of a successful takedown. For example, OS vendors can issue a notice to OS device users notifying the user that the device is vulnerable and needs to be patched or the device contains malware which needs to be removed. Such action will help OS vendors improve the security of their software and improve quality of the service they provide. It is thus in OS vendors’ interests to be aware of any vulnerabilities or malware that is (now or in the future) being exploited by cybercriminals. Consequently, it is in OS vendor' interests to be included in efforts like attempting a botnet takedown.

Suppliers such as ISPs play a central role in DDoS attacks *and* botnet takedowns as DDoS attacks are often executed through ISP networks and it is through ISP networks that botnets can be identified. ISPs can be negatively affected by malicious traffic that is generated by botnets carried through their network because increased traffic from a DDoS attack can make the network conjected and slower for ISP customers (Plohmann, et al., 2011). This means that ISP's customers are directly affected by botnets. Therefore, it is in ISPs interests for botnets to be taken down as they affect the quality of the service provided to customers. As ISPs are a conduit for much of the internet's traffic (van Eeten, et al., 2010), ISPs can take measures if they detect suspicious activity on their network. For example, ISPs can choose to the resolve the problem by restricting online communication or disconnecting an Internet connection altogether (OECD, 2012). Disconnecting a user's internet access, even temporarily, may result in the ISP losing customer business to a competitor as the customer may decide to change providers rather than go through the process of cleaning the infected machine. This may cause issues in terms of loss of business to the ISP as well as issues relating to the termination of a contractual agreement between the ISP and the user if the ISP disconnects a user. Since the United Nations linked internet access to fundamental human rights (OECD, 2012), disconnecting customers could raise legal and commercial/cost-benefit challenges if blocking access to sites is perceived as a form of censorship.[59] Privacy and data protection concerns can arise when collecting and sharing IP addresses of computers that are potentially operating in a botnet. In some countries IP addresses are considered personal information and must be treated appropriately. This means that processing and collecting identifiers like IP addresses and other personal information increases the potential privacy risks for suppliers like ISPs in attempted botnet takedowns under GDPR (European Commission, 2016).

Many countries have their own anti-botnet mitigation policy, programme or centre that is either fully or partially funded by the government or by entirely privately-led initiatives (OECD, 2012). For example, in the Netherlands, ISPs can sign a commitment to notify customers of compromised machines in the form of a botnet treaty. ISPs participating in the anti-botnet 'treaty' are expected to fund their own notification and disinfection activities and will notify users of the problem and isolate their machines as necessary (OECD, 2012). Germany's anti-botnet effort is voluntary and led by the private sector with financial and technical support provided by government. Germany's Federal Ministry funds the technical support services provided to customers whose computers have been

[59] In public health, freedoms can be restricted in circumstances where it is believed that an individual's health or behaviour may negatively impact others. For example, in the case of contagious diseases individual's freedom can be restricted via quarantine.

identified as infected. ISPs notify and support infected customers through various channels such as e-mail, phone or SMS, but they are not permitted to block or quarantine users (a practice known as erecting a 'walled garden') (OECD, 2012). In Australia, ISPs contact customers to inform them that their computer is compromised, explain the potential consequences of not addressing the situation, and inform customers how to fix the machine and how to prevent re-infection. ISPs can decide to reset customers passwords forcing customers to contact the ISP at which point they will be notified of their malware problem. ISPs are allowed to temporarily quarantine an infected computer and/or restrict outbound email messages (by blocking certain network ports) (OECD, 2012).

It is in ISPs' interests to collaborate with organisations to mitigate the effects of botnet takedowns as doing so can result in ISPs providing more secure services to customers, reducing costs associated with technical support and customer service, improving network performance through the management and the reduction of compromised internet connections and strengthening user confidence in ISPs (OECD, 2012). However, there are costs involved and the more involved the ISP is in the takedown, the more costly it is likely to be. For example, deploying various communication channels for notifying customers about infected machines will impose a cost to the ISP. The Irish government for example, fund an anti-botnet programme. In Japan there is a dedicated organisation called the Cyber Clean Centre that assists customers with infected machines. In the United States there is a privately led response and ISPs must cover their own anti-bot efforts including notifying customers of infected machines. Interestingly, where governments cover the costs, anti-botnet policies are more likely to have a higher success rate (OECD, 2012). This is due to disparities between small and large ISPs financial flexibility and resources.

When there is no obligatory programme in place that both protects ISPs from the financial and legal risks associated with botnet takedowns, ISPs who have less financial flexibility are less likely to participate. This suggests that while all ISPs have an interest in improving the security of cyberspace and taking down harmful botnets that spread through their networks, smaller ISPs that are not publicly funded are limited in terms of how involved they can be due to the associated costs that come with notifying customers of infected devices.

The above analysis suggests that an ISP's interest in improving the quality of service provided to their customers is satisfied when an attempted takedown is successful, particularly when the takedown results in deterrence and reduces the prevalence of DDoS attacks on their network. As participation in an attempted takedown does not guarantee success and it can be costly, a failed takedown is not in the interests of ISPs. Depending on the size of the ISP, interests also vary greatly. For example, a small ISP may not have an interest in a successful takedown due to the cost the takedown will incur.

Therefore, this analysis suggests that the interests of ISPs not only hinge on the success or failure of a takedown, they also depend on the potential legal and financial cost of the endeavour. It is in ISPs' interest for governments to fund botnet takedown attempts. ISPs could market their participation in botnet takedowns to governments as an add-on service that adds national value to the cyber environment in which they operate, one that no other service or platforms can provide. Government funded endeavours could provide assurance to governments over ISPs' participation in botnet takedowns, outline national legal requirements and processes to follow, list relevant stakeholders to engage with such as OS vendors, and also provide insurance to ISPs in terms of the cost of the takedown. This could remove the reliance on the success of participation as well as the unknown legal and financial implications.

Local community interests

The local community refers to the area in which the affected organisation is situated. For a physically existing organisation that is a member of a community, the local community thus encompasses individuals and organisations residing within that community. Even online organisations will require a physical address in order to register as a business, regardless of whether the business physically operates from said address or whether the business solely operates online. For example, a clothing drop shipping business is one where products are sourced from a supplier and sent directly from the supplier to the customer. The online store is simply the median connecting the supplier with the customer and making a profit in the meantime. Even though it is not necessary for the focal organisation to have a physical store, they will be required to register the business name and address in order to be recognised as a legal business. In addition, if customers wish to return an item, they will need the e-businesses' forwarding address. In such cases, an e-commerce company's local community will be their suppliers, customers and employees. These stakeholders will have the same interests as those identified in previous section of this chapter and will not be repeated.

For traditional businesses who have a physical residency, the interests of the local community are individuals, businesses and authorities who may or may not be existing or future shareholders, employees, customers and suppliers. The community could be harmed by DDoS attacks that are using bandwidth required for community online services. Community online services could be considered critical such as community health centres which typically list vital information relating to emergency care and services. Or services could be more socially orientated such as sporting events or fund raisers. If websites crash due to the DDoS attack, it is in the local community's interest for the websites to be up and running as soon as possible. It is also in the local community's interest that the cost and collateral damage caused by the initial attack is minimised. A failed botnet takedown that exacerbates the damage and cost of the initial attack through access delays and escalation is not in the local community's interests.

The local community will have a collective interest in taking actions to reduce the likelihood of a similar attack occurring again. It is unlikely that the local community will have the resources to help the affected organisation or protect itself from future attacks. Therefore, it is likely that they will rely on others who are better positioned to take such action on their behalf. This line of thinking suggests that it is in the local community's interest for the appropriate persons or institutions to take measures that will protect those who cannot protect themselves. In respect of an attempted takedown, this interest can be satisfied if a botnet takedown is successful.

Today organisations must be able to withstand cyberattacks as it is not a matter of if an attack will occur, it is a matter of when. The local community can benefit from the long-term success of organisations in their locality. Long-term success can present companies with the opportunity to expand and offer more vacancies to members of the local community. Long-term success will be impacted by the organisations ability to withstand and appropriately respond to cybersecurity attacks like a DDoS attack. Attempting a legally authorised takedown is thus in the local community's interests. A successful collaborative approach that involves engagement with the relevant authorities can ensure the organisation stays within the realms of the law and does not jeopardise the reputation of the community in any way. This is linked to the community's interest in maintaining a good reputation nationally and internationally as a good reputation can either attract or deter future investors. An organisation that invests time and resources in attempting illegal botnet takedowns or continuously invests in failed takedowns that cause harm to local services does not satisfy the local community's interest in attracting future investors and improving the security of cyberspace respectively.

From the above analysis it appears that local community's interests revolve around taking action that positively impacts the community. For example, limiting damage and costs of the initial attack and responding in a legally appropriate way that does not compromise the reputation of the local community, that contributes towards the security of cyberspace and prevents similar attacks from happening in the future. A successful collaborative botnet takedown fulfils these interests.

6.3.2 Derivative stakeholders Interests

In order to stay true to Phillips, this section provides a brief description of derivative stakeholders' interests in relation to the decision to be made. Derivative stakeholders are those who can help or harm the organisation and its normative stakeholders. It is reasonable for management to spend a limited amount of time and resources managing derivative stakeholders' interests (Phillips, 2003b:

222). Stakeholders who can help or harm the organisation in respect of the decision to attempt a botnet takedown include competitors, the media, the DDoS attacker and law enforcement.

Competitors' interests

It is in competitors' interests for their rivals to implement good cybersecurity practices and policies as this can improve cybersecurity posture at an organisational level. Good security practices at organisational level feed into better security within the wider industry community. Adopting good cybersecurity practices and policies is particularly relevant to organisations who operate in sectors that rely on internet-facing applications, for example, gaming and gambling industries.

Competitors have a security and financial interest in reducing the prevalence of malicious attacks in their sector to reduce the volatility of cyberspace and reduce cybersecurity costs. This is a collective interest that competitors share with the organisation. Competitors may be cognizant that they might be just as easily targeted and may see collaboration as an endeavour that is in their interests. It may demonstrate solidarity with the attacked organisation and fortify defences against the attackers. This act of solidarity is likely to depend on the risk of the collective effort being a deterrent versus it being an escalation threat where the latter is unwanted, and the former is desirable. As the relative risk rating is unknown, the decision to collaborate will depend on the chances of the takedown being successful.

DDoS attacks are likely to occur in the public eye e.g., a website's performance is lagging, availability is disrupted as they are externally visible to various stakeholders e.g., the user, network analysts and of course, the attacker. If competitors decide to publicly work in conjunction with the affected organisation to track down the botnet, this may deter the attacker (and possibly other attackers) from launching future attacks out of fear of the collective backlash. A collective agreement between competitors and suppliers may be developed with the aim of lobbying government to participate in funding activities. This will have a positive impact on sharing the costs and risks associated with botnet takedowns which would open opportunities for smaller ISPs to participate more freely in active cyber defences (as mentioned in the supplier's analysis, smaller ISPs are unlikely to engage in botnet takedowns due to costs). This suggests that competitors can gain from collaborating with the affected organisation (and others) based on the assumption that help will be reciprocated should the competitor find themselves in a similar position. If competitors choose to not stand in solidarity with the affected organisation, when they themselves are targeted in a DDoS attack and would like to attempt a botnet takedown, the competitor will be likely to face the costs and risks of doing so alone.

Conversely, competitors can passively gain from their rivals being successfully targeted by malicious outsiders if customers choose to divert their business from the affected organisation to competitors. This can result in more market share for the competitor. Further benefits for competitors emerge when the affected organisation attempts and fails a botnet takedown that causes escalation. While the affected organisation scrambles to manage the escalation, competitors can 'get ahead'. It is thus fair to assume that competitors do not have a primary interest in their rivals falling victim to cybersecurity attacks or responding to cybersecurity attacks that result in escalation; rather competitors have a primary interest in following industry standards so as to best position themselves to respond to a cybersecurity threat like a DDoS attacks.

In rare cases competitors have instigated DDoS attacks targeting their rivals. For example, in 2012 the owner of ChronoPay, a payment service provider, was charged with organising a DDoS attack against a competitor in an attempt to secure a lucrative contract for which the two companies were competing (Krebs, 2011). This is an exceptional case and is not representative of the entire group of competitors. Yet, it reinforces Phillips' strategy to consider competitors as a group of stakeholders who have the power to help or harm an organisation. If competitors could be potential attackers, it is in the interests of the organisation to implement strategies to protect themselves against competitor-driven attacks. For the wider group of competitors i.e. the ones who do not commit cybercrime, it is in their interests for their rivals to stop launching cybersecurity attacks as this is disruptive to cyberspace and the sector within which the affected organisation and competitors operate.

If it becomes clear that a number of organisations from one particular sector are vulnerable to DDoS attacks, this could tarnish the reputation of all members of that industry including competitors. It could also put a mark on that industry ultimately resulting in more attackers targeting specific organisations within the same sector. For attackers, the ease of switching from targeting one organisation to another will be very easy. This suggests that while competitors may gain market-share from a once-off attack on their rivals, a number of attacks on rivals is not in competitors' interests. This points to the above-described benefits of collective action and acting in solidarity.

This analysis suggests that whilst competitors may benefit from rivals falling victim to one DDoS attack and may benefit from rivals delaying their recovery by attempting a botnet takedown and provoking an escalation, it is not in competitors' interests for their rivals to adopt insecure cybersecurity practices that contribute to the volatility in cyberspace. A collaborative successful botnet takedown that results in deterrence seems to satisfy competitor's interest in improving the security of cyberspace. A mutually beneficial agreement between competitors, the affected organisation, suppliers and authorities will present the advantage to competitors of being able share costs and risks associated with takedowns both present and in the future. This is unlikely to come to fruition in the absence of such an agreement. A collaborative takedown that results in escalation does not satisfy competitors interests but again, the costs and risks associated with the failed takedown will be shared.

News Medias' interests

It is in the media's interest to share captivating news stories with their readers/viewers. Whether the story harms or helps the people or companies named in the story is potentially irrelevant (unless the media company have a close affiliation i.e. political, familial or financial with the affected organisation. This means that the media can benefit from producing a botnet takedown story in a favourable or unfavourable light.

An example of a good-news-story could be one that describes a successful, legal takedown. The affected organisation would benefit from such coverage if the active response was described as a collaborative endeavour that benefits a multitude of stakeholders including persons with infected devices. While the story may indicate that the successful takedown helped identify thousands of infected devices, it may also conclude that the investigation resulted in the arrest and prosecution of the perpetrator, which would be advantageous publicity to the affected organisation. Such a story would place the affected organisation in the limelight and may attract more customers. It is also possible that such a story would be a message to all attackers that organisations are successfully collaborating in the fight against cybercrime. The news media may benefit from the latter as it may deter attackers from launching further attacks out of fear of being arrested and prosecuted.

Although, conventional journalistic wisdom suggests that negative news resonates more with readers (a practice commonly referred to by journalists as negativity bias). If negative news stories can evoke stronger psychophysiological reactions e.g., heightened attention and arousal when compared to positive news stories (Ellwood, 2020), it will thus be more beneficial to the media to produce the worst possible outcome for the affected organisation or, at the very least, employ a negative spin of the takedown. For example, the coverage of a failed illegal botnet takedown that resulted in escalation could create a stir amongst viewers/readers. This could be particularly emphatic if the story highlighted the delay in access to products and services because of the provoked escalation as well as the harm caused to innocent third parties that got caught in the crossfire.

The media heavily rely upon the internet, particularly social media platforms such as Facebook, YouTube, Twitter, Instagram and others to share their stories with the public. As there is currently no tool available on the market that provides full protection against DDoS attacks, it is possible that the media may find themselves victim to a DDoS attack. Any disruption to internet access could greatly impede the media's ability to process, disseminate, amend and store valuable information. The media thus have a vested interest in cyberspace becoming less volatile and more secure. It is thus reasonable to assume that the media have an interest in organisations collaboratively working with suppliers and authorities to improve the security of cyberspace. Successfully taking down botnets can contribute towards satisfying this interest. A failed takedown may negatively impact the security of cyberspace through disruption or escalation, and the news media could be on the receiving end of such disruption or escalation. In this case, a failed attempt is thus not in the interests of the news media until they can resume business and write about it.

The above analysis suggests that it is in the media's interest for organisations and authorities to work together in an effort to deter DDoS attackers and reduce the prevalence of DDoS attacks. Successful botnet takedowns may achieve this over time. Until then, it is in the media's interests to write about the most eye-catching interpretation of botnet takedowns, whether that is favourable or unfavourable to the affected organisation.

Attackers' interests

It is important for organisations to stay informed as regards cybersecurity trends, threats and attackers' motivations. This can provide valuable insight into the common vectors attackers exploit. By analysing such information, organisations can craft a defence strategy and mitigation plan. DDoS attacks are typically executed by attackers to cause disruption but can also be used by attackers who have interests in other nefarious activities, for example blackmail as seen in RDDoS attacks. It is in attackers' interests for organisations to be ill-prepared for DDoS attacks in order to reach their goal of disruption or otherwise. It is not in the attacker's interest for the organisation to attempt a botnet takedown as it may disrupt the attacker's network. It may also take down the botnet that the attacker either bought as a service or created themselves.

Law enforcement's interests

In the absence of law enforcement, it is possible that an attempted botnet takedown will be illegal due to the likelihood of bots transcending different jurisdictions. If the organisation engages in illegal activity, law enforcers are likely to investigate. Negative publicity relating to the investigation could harm the organisation. If the affected organisation decides to collaborate with law enforcers to takedown the botnet, this could help the organisation. As such, law enforcers are considered a derivative stakeholder.

The European Union implemented a Directive 2013/40/EU (the 'Botnet Directive') which outlines the law on botnets and their mitigation (European Union, 2013). Directive 2013/40/EU defines five categories of botnet-related crimes which cover illegal access to information systems, illegal system interference, illegal data interference, and the production and sale of computer programs used for committing the named crimes (European Union, 2013). If an organisation were to attempt a botnet takedown it is likely that this would entail data interference and system interference. Doing so without the involvement and prior consent of law enforcement, the takedown is likely to breach the guidelines set out in this directive. It is in law enforcement's interests for organisations to comply with regulations and directives and avoid taking action that may breach them. Otherwise, the organisation may be subject to legal proceedings which is not in the interests of law enforcers or the affected organisation.

A failed botnet takedown may result in a retaliatory attack, which means that two attacks or more have been perpetuated by how the affected organisation chose to respond. As alluded to in section 6.2.3, cybercrime is growing at a rate that law enforcers are unable to manage or control, and arrest and prosecution rates are extremely low. It is not in law enforcers' interests for organisations to fail in their attempts and perpetuate the problem by provoking more cybersecurity incidents. It is in law enforcements' interest to collaborate with organisations to help the process of tracking and tracing attackers in a legal and more controlled environment. If collaboration is successful, this will satisfy law enforcer's interest in mitigating the damage caused by the initial attack. Taking down the source of the original attack prevents a second attack emanating from the same source. Working in conjunction with the affected organisation and suppliers will also better position authorities to create a case against the attackers and use in prosecutions.

6.3.3 Prioritisation of conflicting stakeholders' interests

The analysis of stakeholders' interests suggests that conflict does not arise between the interests of different normative stakeholders. This means that the prioritisation of stakeholders' interests is not necessary as all stakeholders appear to have similarly placed interests in respect of the attempted takedown being a success or failure.

6.4 Conclusion

Normative stakeholders have a collective interest in the attempted takedown being a success. The likelihood of success (which includes deterrence, and sharing the cost and risk associated with the attempted takedown) is very likely to be strengthened by collaboration with authorities and suppliers. However, there is no risk rating available in terms of the probability of success versus the probability of failure. For the sake of thoroughness, let us consider the likelihood of a successful takedown versus the likelihood of failure.

The first case presented is a successful takedown i.e., the best case scenario. An ideal takedown would be one that does not cause harm to third parties yet successfully disables the botmaster's system. One that does not incite a retaliatory attack or escalation, but rather deterrence. One where the botnet is legally taken down, the owners of the compromised devices are notified and the devices are cleansed. One that does not result in additional financial losses for ISPs, privacy infringing legal proceedings or further delays to employee access or customer services. One that involves the identification of the attacker and prosecution. A successful takedown will require involvement with authorities and suppliers to ensure the interests of all normative stakeholders.

The polar opposite to the ideal attempt is an unsuccessful botnet takedown i.e., the worst-case scenario. One where the victim organisation uses corporate time and resources in a failed attempt at a botnet takedown. One that may involve an illegal or legal takedown that extends downtime, delays normal working conditions and access to customer services. One that either reinvigorates the attacker to update the botnet making it more resilient to a takedown or incites the attacker to launch a retaliatory attack causing more damage than the initial assault. One where the attacker is not deterred and there may be a further financial cost to ISPs who lose business to competitors. This worst-case scenario is not in the interests of all normative stakeholders.

The first apparent problem with the best-case scenario presented is that many elements are unlikely to occur or are in fact, unattainable. For example, harm to third parties cannot be avoided as all the available technologies for tracing an internet attack to its source must go through the zombies used to stage it and hence inevitably trespass onto innocent machines. This means that any active defence strategy remotely likely to succeed in stopping a sophisticated attack will invariably impact the property of innocent persons in some way (Himma, 2004). Furthermore, based on current prosecution rates it is unlikely that the ideal scenario is achievable as it involves identifying the attacker, arresting, prosecuting and deterring them from launching further attacks. A collaborative approach that involves law enforcement will make the process of gleaning evidence required to make against the attacker easier. However, prosecution is not guaranteed nor is deterrence.

An additional problem with the ideal scenario relates to the legality of the takedown. Botnets are invariably spread across more than one legal jurisdiction or market. The legal right to privacy can vary depending on the applicable legal framework. This raises questions relating to how infected machines are identified and with whom information such as a customer's IP address and account details is shared (OECD, 2012). Maintaining the confidentiality of information that is shared across international borders is of particular importance to the organisation and ISP as compromising private information could result in legal proceedings and negatively impact corporate image. Furthermore, it is paramount that the organisation and ISP act legally in order to ensure protection against compromised device owners if the device owners decide to take legal action over the harm caused to them as a result of the takedown. At present, there is an absence of adequate state funding that offers private communication between international bodies, ISPs and organisations. There is also a lack of clear policies and protections for those executing the takedown, i.e. private organisations and ISPs. It would therefore be remiss for the victim organisation and ISP to engage in a botnet takedown without the involvement of law enforcement as this could place them in a compromising legal position, one that may jeopardise their reputation and business. This again, points to the added benefit of involving law enforcement to properly assess the potential legal ramifications of engaging in a takedown.

In the ideal scenario, it is assumed that the ISP is willing to be actively involved in the botnet takedown. However, this is not guaranteed. As mentioned previously in this chapter, smaller ISPs may be unwilling to voluntarily commit to engaging in a botnet takedown due to operational costs. For example, the ISP may not have the budget or infrastructure to inform or offer support to the compromised device owners.

By juxtaposing the best and worst case scenarios it is clear that it is unlikely that the best-case scenario is possible at the time of writing. If the best-case scenario is unlikely to occur, this means that the organisation would be investing corporate time and resources in an endeavour that may provide some benefit but there is no guarantee the endeavour will provide any favourable outcome for the organisation and its stakeholders. It is thus reasonable to assume that the best-case scenario does not, at present, support the continuation of the cooperative scheme and is thus, unlikely to reach the communicative assent of all normative stakeholders.

In the worst-case scenario, it appears that attempting a botnet takedown may incite a further DDoS attack, cause more disruption to cyberspace, supplier and customer services, harm innocent bystanders and business operations. It is fair to assume that any one of these unfavourable outcomes is a likely possibility. This suggests that the worst-case scenario does not support the continuation of the cooperative scheme and is unlikely to achieve the communicative assent of all stakeholders.

It is clear that attempting a botnet takedown only serves the interests of normative stakeholders when the outcome is a success and the probability of success is strengthened by a collaborative approach. If management were able to determine that the chances and benefits of success outweigh the risks/costs, an attempted takedown would support the continuation of the cooperative scheme. Management would also be able to communicate this to normative stakeholders which is then likely to reach the communicative assent of all stakeholders. Given that management are unlikely to be in a position to determine that the chances and benefits of success outweigh the risks/costs, an attempted botnet takedown is unlikely to reach the communicative assent of all stakeholders, or support the continuation of the cooperative scheme. In the absence of a risk rating, this analysis concludes that is it not ethically appropriate for organisations to attempt a botnet takedown.

Chapter 7 Conclusion

7.1 Summary of the results

The results are presented alongside the following four research questions analysed in this thesis:

1) What are the ethical issues that arise in cybersecurity in the business domain? Are there any blind spots in the ethical literature that are worthy of further ethical deliberation?
2) Is it ethically appropriate for organisations to employ red teams to find security vulnerabilities?
3) What is the ethically appropriate organisational response to a ransomware attack?
4) Is it ethically appropriate for organisations to attempt a botnet takedown in response to a DDoS attack?

7.1.1 What are the ethical issues that arise in cybersecurity in the business domain? Are there any blind spots in the ethical literature that are worthy of further ethical deliberation?

Chapter 2 highlights that various ethical issues arise in cybersecurity in the business domain (see section 2.3.2, Table 1 for more details). It shows that ethical issues do not always arise in the same context. For example, privacy is raised as an ethical issue in cases where organisations are secretly monitoring employees in the absence of consent. It also raised as an ethical issue in relation to organisations sharing personal information with third parties in the absence of informed consent.

Chapter 2 also pinpoints three blind spots (ethical hacking, responding to cybersecurity attacks, the IoTs) in the literature as important cybersecurity topics that require further ethical scrutiny (see section 2.3.2 for more details). These blind spots were chosen for further analysis; ethical hacking (employing red teams), organisational responses to cybersecurity attacks (responding to ransomware and Distributed Denial of Service (DDoS) attacks) and ubiquitous devices (DDoS attacks are enabled by the widespread use of ubiquitous devices i.e., devices connected to the internet also known as the Internet of Things (IoTs).

All three blind spots are not only relevant to organisations, but they also interconnect with one another. For example, DDoS attacks and ransomware attacks are growing cybersecurity threats to organisations. Ethical hackers can aid the management of cybersecurity threats by running tests that identify vulnerabilities in people and technology. For example, a social engineering test can determine how vulnerable employees are to clicking a phishing email which, in real life, could potentially launch a ransomware attack. Ransomware attacks are also being used by attackers in conjunction with DDoS attacks (referred to as RDDoS). This typically happens when attackers do not receive a ransom in a timely fashion and launch a DDoS attack to send a message to the victim to expedite payment. DDoS attacks are not possible to execute without the existence of insecure devices connected to the internet i.e., IoTs. These blind spots are thus relevant to 1) organisations that are being targeted or may potentially be targeted by DDoS and ransomware attacks 2) the field of ethics as stakeholders' interests may be affected by how organisations choose to respond and mitigate these threats e.g., by employing red teams and 3) cybersecurity as the goal of cybersecurity is to protect devices and people who have access to them.

7.1.2 Is it ethically appropriate for organisations to employ red teams to find security vulnerabilities?

Chapter 4 highlights that employing red teams is a gold standard approach that can be utilised by management to improve the security posture of an organisation. Yet, red teaming involves manipulating people and technology in order to discover whether any existing vulnerabilities could be exploited by real attackers. The goal of the red team is to imitate a real-attacker and mimic tactics employed by malicious attackers. Employees are thus unaware of red team tactics and testing as this allows the red-team to test the competency of employees through social engineering attempts (a tactic successfully used by malicious attackers to gain access to infrastructure, devices, files etc). Given that employees are unaware of testing and are targeted by a red team but are normative stakeholders whose interests must be considered by management in the decision making process, red teaming thus raises the ethical question, is it ethically appropriate for red teams to manipulate people and technology to improve the security posture of organisations? Chapter 4 contains an ethical analysis of this pertinent question.

The analysis suggests that it is ethically appropriate for organisations to employ a red team in an attempt to reduce the margin of security vulnerabilities created or facilitated by people and technology. Red team testing can reveal weak spots in an organisation's infrastructure which can be rectified before those weaknesses are exploited by adversaries. Due to the esteemed expertise of red teams and the large scope of testing that can be undertaken, it is paramount that testing is limited in order to reduce the negative impact an engagement may have on the organisations and its key stakeholders.

Four recommendations are provided in Chapter 4 which enable red teamers to test for vulnerabilities in people and technology in a way that will reduce the likelihood of the engagement negatively impacting the sponsor organisation and its key stakeholders. Those recommendations are: 1) the engagement is limited to the extent that it simulates real-world attacks in a fashion that does not cause unnecessary harm to people or technology, 2) the constraints of the rules of engagement do not negate the purpose of a red team engagement which is to find vulnerabilities in people and technology,

3) red team reporting includes listing the names of employees who contributed towards the success of the simulated attack for education and training purposes for the named-employees only, and 4) the red team have the ability to and are willing to constructively share the results of testing with the internal security team.

7.1.3 What is the ethically appropriate organisational response to a ransomware attack?

At the time of submission (June 2021), new stories permeated the globe surrounding a ransomware that successfully targeted the Health Service Executive (HSE) in Ireland (Halpin & Humphries, 2021; Mehta, 2021; MacNamee, 2021; Perlroth, 2021). The HSE attack forced the HSE to shut down all IT systems to reduce collateral damage caused by the attack (Halpin & Humphries, 2021). In the absence of IT, healthcare workers are being forced to use paper records to keep services operational which is causing severe disruption to services during the COVID-19 pandemic (Perlroth, 2021). €16.3million was demanded by the attackers in return for access and the HSE publicly stated they will not pay the ransom (MacNamee, 2021). This case and the entirety of Chapter 5 highlight ransomware as a growing threat to organisations, the serious impact a ransomware attack can have on organisations and the extensive costs that can incur in respect of downtime and recovery. Furthermore, every organization from any industry can fall victim to a ransomware attack as no organization is immune. To date those who have fallen victim to ransomware attacks have individually chosen their own path in respect of the decision to pay, not pay or negotiate with the attackers.

The analysis in Chapter 5 is a critique of Phillips' approach. Phillips' method does not differentiate between short- and long-term interests. The author believed in the case of responding to a ransomware attack it was necessary to consider stakeholders' short- and long-term interests in isolation based on the likelihood that the managerial decision to be made was likely to affect stakeholders short- or long-term interests differently. This critique was applied in Section 5.3.1 and the results of the example analysis confirm that short- and long-term interests vary.

The analysis undertaken indicates that the decision to pay, not pay or negotiate is not straightforward, in fact arriving at a decision can be extremely complicated. It appears that the ethically appropriate response to a ransomware attack depends on the affected organisation and the context of the decision to be made (see Chapter 5 for more details). This means that despite the example analysis suggesting that not paying the ransom is the ethically appropriate choice (see Section 5.3 for more details), this may not *always* be the case. As explained in Sections 5.3 and 5.4, if not paying the ransom expedites the demise of the targeted organisation, not paying would not support the continuation of the cooperative scheme. Therefore, paying or negotiating a lower ransom could be deemed as the appropriate ethical response. This means that ransomware cases need to be analysed on a case-by-case basis to determine which response is ethically appropriate.

7.1.4 Is it ethically appropriate for organisations to attempt a botnet takedown in response to a DDoS attack?

Chapter 6 is an acknowledgement of the staggering increase in DDoS attacks and the lack of appropriate responses available to organisations. This chapter explains how the proliferation of insecure devices connected to the internet e.g. cameras, smart TVs, radios, printers, can be easily manipulated by malicious hackers to create a network of bots called a botnet. Botnets are then used to launch DDoS attacks which can cause serious disruption to the target organisation. A great example of the disruption a DDoS attack can cause is the Mirai botnet which comprised of 100,000 infected devices effecting the operations at many large MNCS including Deutsche Telekom (taking 900,000 of their customers offline), Airbnb, Netflix, PayPal, Visa, Amazon, The New York Times, Reddit and GitHub. (Russell, 2017).

DDoS attacks also continue to rise in frequency and size due to the COVID-19 pandemic, the widespread availability of DDoS-as-a-service and attackers becoming more sophisticated in their approach i.e. using attack vectors in conjunction with DDoS attacks such as RDDoS (IDG, 2018). The COVID-19 pandemic forced organisations to embrace remote working, something which organisations were unprepared to support. Organisations security posture weakened due to changes in work and infrastructure patterns (ENISA, 2020b) and security risks have increased as a result of more home internet usage. For example, one security company, Lumen reported a monthly increase of 1200% in emergency DDoS mitigation activations between the months of July and October in 2020 (Lumen, 2020).

Organisations can choose how to respond to DDoS attacks as they have the option to defend and reduce collateral damage or they can try to attempt to takedown the botnet through which the attack was launched. The latter presents ethical questions as a botnet takedown can result in causing harm to third parties. Botnet takedowns could also break the law as botnets can pervade many jurisdictions resulting in potential legal proceedings for the organisation. The analysis in Chapter 6 thus focuses on the ethics of a botnt takedown in an attempt to determine whether it is ethically appropriate to takedown a botnet in response to a DDoS attack.

The ethical analysis in Chapter 6 recommends that it is not ethically appropriate for organisations to attempt a botnet takedown in response to a DDoS attack unless 1) a collaborative approach is taken and 2) it is possible to determine that the likelihood of deterrence outweighs the risk of escalation. The analysis suggests that a collaborative approach will require mandatory involvement of law enforcement and ISPs. Involvement is described as mandatory as both parties increase the likelihood of success and deterrence. For example, law enforcers can provide advice on the legal risks of the attempted takedown while ISPs are best positioned to disconnect infected bots running through their network and may be able to identify the location of the C&C server. ISP involvement will hinge on government funding due to the potential costs that attempting a botnet takedown will incur. This is especially relevant to smaller ISPs who have limited resources.

7.2 Limitations

The study at hand has the following three limitations.

1. Resources for this thesis were limited to the English language only (see Section 2.2 for more details) and within a specific timeframe (from 1996 onwards). This is a limitation of this research as informative publications may exist in languages other than English or in resources published prior to 1996.
2. It is possible that organisations who offer red teaming as-a-service have developed their own set of in-house ethical rules or guidelines that are not readily available for public or academic consumption. If in-house ethical rules or guidelines exist, they could provide valuable insight into a red teaming engagement. It is equally possible that no in-house ethical rules or guidelines exist in commissioned organisations. Either way, not having attempted to gain access to those rules or guidelines is a limitation of the study at hand.
3. Organisations who manufacture and distribute IoTs devices may follow in-house security-by-design guidelines that are not readily available for public or academic consumption. If such guidelines exist, they may outline the importance of implementing security pre-production. The guidelines may outline with whom responsibility and accountability lies pre-production and the potential impact of releasing insecure devices to market. It is equally possible that no in-house ethical guidelines exist in organisations that manufacture IoTs devices. Not having access to such rules or guidelines (if they exist) is a limitation of this research.

7.3 Future research

This thesis presents three potentially interesting avenues of future research. They are:

1. Future research that includes resources from languages other than English and sources pe-1996 might reveal new insights.
2. It would be prudent to undertake empirical research to ascertain whether organisations that offer red teaming as-a-service follow any in-house ethical guidelines that are not available for public or

academic consumption. Such a study would require direct engagement with a number of IS organisations and would be worthwhile as it may reveal for example, that the majority of IS organisations surveyed do not follow any in-house ethical standards, or that the majority of IS organisations surveyed have in-house ethical standards in writing, but they are never followed in practice, lack depth, lack practicality and/or require fine tuning. Such revelations would reinforce the significance of the recommendations listed in Chapter 4 as the recommendations are designed to guide both the sponsor organisation and the commissioned red team throughout a red team engagement.

If the empirical study reveals that some IS organisations utilize existing in-house guidelines, it is likely that the rules or guidelines developed by the IS organisation were designed from an IS point of view only, as opposed to guiding both the sponsor organisation and the commissioned red team. As a result, the IS guidelines may or may not be missing key interests held by the sponsor organisation and its stakeholders. If useful in-house guidelines are identified in the empirical study, it could be worth amalgamating them with the recommendations provided in Chapter 4. In combination, they could be a steppingstone towards developing a universal standard for sponsor organisations and red teams to follow.

3. Manufacturers are failing to implement adequate security to IoTs devices pre-production. This is perpetuating the issue of insecure devices being manipulated by adversaries to launch large-scale DDoS attacks. An interesting area to research could involve engaging directly with manufacturers of IoTs devices to establish whether security-by-design guidelines exist. If they do exist, it is worth investigating the efficacy of those guidelines to determine how they could be improved or better implemented. If ethical guidelines do not exist, the contents and results from Chapter 6 could be used as a starting point to highlight the enabling role manufacturers can play in DDoS attacks. If specific ethical requirements were developed and implemented, this could reduce the number of insecure IoTs devices going to market. As a result, the number of vulnerable IoTs devices may decrease which could reduce the frequency of successful DDoS attacks.

Bibliography

Aasmann, L., 2011. *Legal Aspects of Fighting Botnets - The Estonian Perspective.* [Online]
Available at: https://www.enisa.europa.eu/events/botnets/workshop-presentations/lauri-aasmann-presentation/
[Accessed 24 January 2021].

Abhishta, A., 2019. *The Blind Man and the Elephant: Measuring Economic Impacts of DDoS Attacks,* Twente: Universiteit Twente.

Abrams, L., 2020. *Ransomware gangs and DDoS attacks to their extortion arsenal.* [Online]
Available at: https://www.bleepingcomputer.com/news/security/ransomware-gangs-add-ddos-attacks-to-their-extortion-arsenal/
[Accessed 18 October 2020].

Abreu, R., David, F. & Segura, L., 2016. E-banking services: Why fraud is important? *Journal of Information Systems Engineering & management.* 1(2), p. 120.

Abreu, R. et al., 2015. *Ethics and Fraud in E-Banking Services.* New York, IEEE.

Accenture Security, 2020. *Third Annual State of Cyber Resilience: Lessons from Leaders to Master Cybersecurity Execution,* Dublin: Accenture Security.

Accenture, 2019. *The Ninth Annual Cost of Cybercrime Study: Unlocking the value of improved cybersecurity protection,* s.l.: s.n.

ACM, 2018. *ACM Code of Ethics and Professional Conduct.* [Online]
Available at: https://www.acm.org/code-of-ethics
[Accessed 20 July 2020].

Acunetix, 2020. *Google Hacking: What is a Google Hack?* [Online]
Available at: https://www.acunetix.com/websitesecurity/google-hacking/
[Accessed 10 June 2020].

Akey, P., Lewellen, S. & Liskovich, I., 2018. *Hacking Corporate Reputations.* [Online]
Available at: https://ssrn.com/abstract=3143740
[Accessed 18 June 2020].

Alouane, M. & Bakkali, H. E., 2015. *Security, Privacy and Trust in Cloud Computing: A Comparative Study..* s.l., s.n., pp. 1-8.

Altham, J., 1973. Rawls's Difference Principle. *Philosophy,* 48(183), p. 75.

Angrishi, K., 2017. *Turning Internet of Things (IoT) into Internet of Vulnerabilities (IoV): IoT Botnets.* [Online]
Available at: https://www.semanticscholar.org/paper/Turning-Internet-of-Things(IoT)-into-Internet-of-%3A-Angrishi/08f4b3d292d801878ec3b0445e35ef4ad8c5193e
[Accessed 24 September 2020].

AppRiver, 2019. *Cyber Threat Index for Business Survey.* [Online]
Available at: https://www.appriver.com/files/documents/cyberthreatindex/AppRiver-Cyberthreat-

Index-for-Business-Survey-exec-summary_FINAL.pdf
[Accessed 18 July 2019].

Arghire, I., 2020. *TrickBot Gets Updated to Survive Takedown Attempts.* [Online]
Available at: https://www.securityweek.com/trickbot-gets-updated-survive-takedown-attempts
[Accessed 18 January 2021].

Armerding, T., 2018. *The 17 Biggest Data Breaches of the 21st Century: Security Practitioners weigh in on the 17 worst data breaches in recent memory.* [Online]
Available at: https://www.csoonline.com/article/2130877/data-breach/the-big
[Accessed 18 June 2018].

Atkinson, S., 2018. *Cybersecurity Tech Basics: Vulnerability Management: Overview.* [Online]
Available at: https://www.cisecurity.org/wp-content/uploads/2018/07/Cybersecurity-Tech-Basics-Vulnerability-Management-Overview.pdf
[Accessed 4 May 2020].

Aurangzeb, S., Aleem, M., Iqbal, A. M. & Islam, A. M., 2017. Ransomware : A Survey and Trends. *Hournal of Information Assurance and Security,* Volume 12.

Axelrod, R. & Iliev, R., 2014. Timing of cyber conflict. *PNAS,* 111(4), p. 1–6.

Baha, A. & Luppicini, R., 2016. echnoethical inquiry into ethical hacking at a Canadian University.. *Int J Techn,* 7(1), p. 62–76.

BCS, 2017. *23 Years of DDoS Attacks.* [Online]
Available at: https://www.bcs.org/content-hub/25-years-of-ddos-attacks/
[Accessed 17 September 2020].

Beazley, 2019. *https://www.beazley.com/news/2019/beazley_breach_briefing_2019.html,* London: Beazley.

Bennasar, H., Essaaidi, M., Bendahmane, A. & Benothman, J., 2015. *In Proceedings of 2015 Third IEEE World Conference on Complex Systems (WCCS).* Marrakech, Morocco, IEEE.

Berknik, I. & Mesko, G., 2012. *Study of the Perception of Cyber Threats and the Fear of Cybercrime. ICIW2012.* s.l., s.n., p. 28.

Bianchi, D. & Tosun, O. K., 2019. *Cyber Attacks and Stock Market Activity.* [Online]
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3190454
[Accessed 16 June].

Bisson, D., 2015. *DD4BC Group Targets Companies with Rnasom-Driven DDoS Attacks.* [Online]
Available at: https://www.tripwire.com/state-of-security/security-data-protection/cyber-security/dd4bc-group-targets-companies-with-ransom-driven-ddos-attacks/
[Accessed 22 January 2021].

Bodle, R., 2011. Regimes Of Sharing: Open APIs, Interoperability, and Facebook. *Information Communication & Society,* 14(3), p. 320–37.

Bonner, S. & O’Higgins, E., 2010. Music Piracy: Ethical Perspectives. *Management Decision,* 48(9), pp. 1341-54.

Brangetto, P., Caliskan, E. & Roigas, H., 2015. *Cyber Red Teaming: Organisational, technical and legal implications in a military context.* [Online]
Available at: https://ccdcoe.org/uploads/2018/10/Cyber_Red_Team.pdf
[Accessed 22 July 2020].

Brenner, J., 2014. Information: A Personal Synthesis. *Information,* Volume 5, p. 134–70.

Brey, P., 2007. Ethical Aspects of Information Security and Privacy. In: *In Security, Privacy, and Trust in Modern Data Management Data-Centric Systems and Applications.* Berlin/Heidelberg: Springer , pp. 21-36.

Brink, D., 2007. *Stanford Encyclopaedia of Philosophy.* [Online]
Available at: https://plato.stanford.edu/entries/mill-moral-political/
[Accessed 20 September 2019].

Bromium, 2016. *The hidden costs of detect-to-protect security.* [Online]
Available at: https://learn.bromium.com/rprt- hidden-costs.html.
[Accessed 7 July 2019].

Burt, T., 2020. *New action to combat ransomware ahead of U.S. elections.* [Online]
Available at: https://blogs.microsoft.com/on-the-issues/2020/10/12/trickbot-ransomware-cyberthreat-us-elections/
[Accessed 18 January 2021].

Calyptix Security, 2017. *Biggest Cyber Attacks 2017: How They Happened.* [Online]
Available at: https://www.calyptix.com/top-threats/biggest-cyber-attacks-2017-happened/
[Accessed 20 June 2018].

Case C-222/04 Ministero dell'Economia e delle Finanze v Cassa di Risparmio di Firenze SpA and Others (2004).

Cavusoglu, H., Mishra, B. & and Raghunathan, S., 2004. The effect of internet security breach announcements on market value: Capital market reactions for breached firms and internet security developers. *International Journal of Electronic Commerce,* Volume 9, p. 70–104.

CGI, 2017. *White Paper: The Cyber-Value Connection.* [Online]
Available at: https://www.cgi-group.co.uk/en-gb/white-paper/cyber-value-connection
[Accessed 29 November 2019].

Chang, W., Mohisen, A., Wang, A. & Chen, S., 2015. *Measuring Botnets in the Wild: Some New Trends.* New York, ACM , pp. 645-650.

Chatham House, 2020. *Catham House Rule.* [Online]
Available at: https://www.chathamhouse.org/chatham-house-rule
[Accessed 14 April 2020].

Choi, K., Scott, T. & LeClair, D., 2016. Ransomware Against Police: Diagnosis of Risk Factors via Application of Cyber-Routine Activities Theory. *International Journal of Forensic Science & Pathology,* 4(7), pp. 253-258.

Christen, C. et al., 2014. *Empirically Informed Ethics. Morality between Facts and Norms.*. Berlin: Springer.

Christen, M., Gordijn, B. & Loi, M., 2020. *The Ethics of Cybersecurity.* 1st ed. Cham: Springer Nature Switzerland AG.

Cimpanu, C., 2019. *Norsk Hydro will not pay ransom demand and will restore from back ups.* [Online]
Available at: https://www.google.com/amp/s/www.zdnet.com/google-amp/article/norsk-hydro-will-not-pay-ransom-demand-and-will-restore-from-backups/
[Accessed October 2019].

Cloudfare, 2021. *DNS Amplification Attack.* [Online]
Available at: https://www.cloudflare.com/en-au/learning/ddos/dns-amplification-ddos-attack/
[Accessed 9 May 2021].

CloudFlare, 2020b. *SYN Flood Attack.* [Online]
Available at: https://www.cloudflare.com/learning/ddos/syn-flood-ddos-attack/
[Accessed 6 November 2020].

CloudFlare, 2020. *Famous DDoS attacks: The largest DDoS attacks of all time.* [Online]
Available at: https://www.cloudflare.com/learning/ddos/famous-ddos-attacks/
[Accessed 5 October 2020].

Cole, E., 2001. *Hackers Beware: Defending your network from the wily hacker.* 1st ed. Indiana: New Riders Publishing.

Conger, S., Pratt, J. & Loch, K., 2013. Personal Information Privacy and Emerging Technologies.. *Information Systems Journal ,* September , 23(5), p. 401–17.

Conrad, E., Misenar, S. & Feldman, J., 2010. *CISSP Study Guide.* s.l.:Syngress.

Consultancy.eu, 2020. *Number of claims on cyber insurance policies rising steeply.* [Online]
Available at: https://www.consultancy.eu/news/4695/number-of-claims-on-cyber-insurance-policies-rising-steeply
[Accessed 14 January 2021].

Coveware, 2019a. *Ransom amounts rise 90% in Q1 as Ryuk increases.* [Online]
Available at: https://www.coveware.com/blog/2019/4/15/ransom-amounts-rise-90-in-q1-as-ryuk-ransomware-increases
[Accessed 16 July 2019].

Coveware, 2019b. *Ransomware Amounts Rise 3x in Q2 as Ryuk & Sodinokibi Spread.* [Online]
Available at: https://www.coveware.com/blog/2019/7/15/ransomware-amounts-rise-3x-in-q2-as-ryuk-amp-sodinokibi-spread
[Accessed 24 October 2019].

Coveware, 2019c. *Ransomware Costs Double in Q4 as Ryuk, Sodinokibi Proliferate.* [Online]
Available at: https://www.coveware.com/blog/2020/1/22/ransomware-costs-double-in-q4-as-ryuk-sodinokibi-proliferate

Coveware, 2021. *Ransomware Attack Vectors Shift as New Software Vulnerability Exploits Abound.* [Online]
Available at: https://www.coveware.com/blog/ransomware-attack-vectors-shift-as-new-software-vulnerability-exploits-abound
[Accessed 27 May 2021].

Crane, S., Larsen, P., Brunthaler, S. & Franz, M., 2013. *Booby Trapping Software.* Banff, Canada: NSPW.

CREST, 2019. *CODE OF ETHICS: For Suppliers of Cyber Security Services.* [Online]
Available at: https://www.crest-approved.org/wp-content/uploads/Code-of-Ethics-v3.pdf
[Accessed 20 July 2020].

Da Veiga, A., 2016. *A Cybersecurity Culture Research Philosophy and Approach to Develop a Valid and Reliable Measuring Instrument.* London, UK: IEEE, p. 1006–15.

De Groot, J., 2019. *A History of Ransomware Attacks: The Biggest and Worst Ransomware Attacks.* [Online]
Available at: https://digitalguardian.com/blog/history-ransomware-attacks-biggest-and-worst-ransomware-attacks-all-time#4
[Accessed 4 November 2019].

Dean, M., Payne, D. & Landry, B., 2016. Data Mining: An Ethical Baseline for Online Privacy Policies. *Journal of Enterprise Information Management,* 29(4), p. 482–504.

DeMarco, J., 2018. An approach to minimising legal and reputational risk in Red Team hacking exercises. *Computer Law & Security Review,* Volume 34, pp. 908-911.

Denning, D. E., 2009. The Ethics of Cyber Conflict.. *The Handbook of Information and Computer Ethics,* p. 407–428.

Denning, D. & Strawser, B., 2017. Active Cyber Defense: Applying Air Defense to the Cyber Domain. In: G. Perkovich & A. Levite, eds. *Understanding Cyber Conflict.* Washington, DC: Georgetown University Press, pp. 193-210.

Department of Homeland Security, 2017. *Alert (TA17-181A).* [Online]
Available at: https://www.us-cert.gov/ncas/alerts/TA17-181A
[Accessed 4 November 2019].

Dhillon, G., Oliveira, T., Susarapu, S. & Caldeira, M., 2016. Deciding between Information Security and Usability: Developing Value Based Objectives. *Computers in Human Behavior,* August , Volume 61, pp. 636-66.

D’Arcy, J. & Hovav, A., 2009. Does One Size Fit All? Examining the Differential Effects of IS Security Countermeasures. *Journal of Business Ethics,* Issue 89, p. 59–71.

Dittrich, D., 2015. *Taking Down Botnets - Background.* [Online]
Available at: https://www.iab.org/wp-content/IAB-uploads/2015/04/CARIS_2015_submission_21.pdf
[Accessed 19 January 2021].

Dittrich, D. & Himma, K. E., 2006. Active Response to Computer Intrusions. In: H. Bidgoli, ed. *Handbook of Information Security, Threats, Vulnerabilities, Prevention, Detection and Management: Volume 3.* New Jersey: John Wiley & Sons, Inc., pp. 664-681.

Dodig-Crnkovic, G., 2004. *On the Importance of Teaching Professional Ethics to Computer Science Students.* London, College Publications.

Edgescan, 2019. *Edgescan Vulnerability Stats Report 2019.* [Online]
Available at: https://www.edgescan.com/wp-content/uploads/2019/02/edgescan-Vulnerability-Stats-Report-2019.pdf
[Accessed 14 April 2020].

Ellwood, B., 2020. *Negative news evokes stronger psychophysiological reactions than positive news.* [Online]
Available at: https://www.psypost.org/2020/03/negative-news-evokes-stronger-psychophysiological-reactions-than-positive-news-56180
[Accessed 15 May 2021].

ENISA, 2016. *The Cost of incidents affecting CIIs. Systematic review of studies concerning the economic impact of cyber-security incidents on critical information infrastructures (CII).,* s.l.: ENISA.

ENISA, 2017. *Overview of cybersecurity and related terminology.* [Online]
Available at: https://www.enisa.europa.eu/publications/enisa-position-papers-and-opinions/enisa-overview-of-cybersecurity-and-related-termino
[Accessed June 2018].

ENISA, 2018 May 15. *Cyber Security Breaches Survey 2018.* [Online]
Available at: https://www.enisa.europa.eu/news/member-states/cyber-security-breaches-survey-2018
[Accessed June 2018].

ENISA, 2020a. *ENISA ETL Report 2020: From January 2019 - April 2020: Distributed denial of service,* Marousi: ENISA.

ENISA, 2020b. *ENISA Threat Landscape 2020: Cyber Attacks, Becoming More Sophisticated, Targeted, Widespread and Undetected.* [Online]
Available at: https://www.enisa.europa.eu/news/enisa-news/enisa-threat-landscape-2020
[Accessed 6 November 2020].

ENISA, 2021a. *Botnets.* [Online]
Available at: https://www.enisa.europa.eu/topics/csirts-in-europe/glossary/botnets
[Accessed 3 April 2021].

ENISA, 2021b. *DNS Sinkhole.* [Online]
Available at: https://www.enisa.europa.eu/topics/csirts-in-europe/glossary/dns-sinkhole
[Accessed 8 May 2021].

ENISA, 2021c. *Malware.* [Online]
Available at: https://www.enisa.europa.eu/topics/csirts-in-europe/glossary/malware
[Accessed 3 April 2021].

Ernst Young, 2019. *CEO Imperative Global Challenges.* [Online]
Available at: https://www.ey.com/en_gl/growth/ceo-imperative-global-challenges
[Accessed 5 February 2020].

EU Comission, 2012. *PRESCIENT: Deliverable 3: Privacy, data protection and ethical issues in new and emerging technologies: Assessing citizens' concerns and knowledge of stored personal data.* Brussels: EU Comission.

EU Commision , 2013. *Cybersecurity Strategy of the European Union: An Open, Safe and Secure Cyberspace.* Brussels: EU Commision.

EU Commision, 2008. *Data Protection in the European Union: Citizens' perceptions.* Brussels: European Commision.

EU Commision, 2008. *PASR – Preparatory Action on the enhancement of the European industrial potential in the field of Security research.* Brussels: EU Commision.

EU Commision, 2008. *PRISE D5.8 Synthesis Report – Interview Meetings on Security Technology and Privacy.* Brussels: EU Commision.

EU Commision, 2010. *SPECIAL EUROBAROMETER 359 Attitudes on Data Protection and Electronic Identity in the European Union REPORT Publication.,* Brussels: European Union.

EU Commision, 2012. *2012 Special Eurobarometer 390 on Cybersecurity.* Brussels: EU Commision.

EU Commision, 2015. *PRISMS: The Privacy and Security Mirrors: Towards a European framework for integrated decision making. Deliverable 10.1: Report on statistical analysis of the PRISMS survey.* Brussels: EU Commision.

EU Commision, 2015. *SURPRISE: Surveillance, Privacy and Security: A large scale participatory assessment of criteria and factors determining acceptability and acceptance of security technologies on Europe. D 6.10,* Brussels: s.n.

European Central Bank, 2018a. *TIBER-EU FRAMEWORK: How to implement the European framework for Threat Intelligence-based Ethical Red Teaming.* [Online]
Available at: https://www.ecb.europa.eu/pub/pdf/other/ecb.tiber_eu_framework.en.pdf
[Accessed 20 July 2020].

European Central Bank, 2018b. *What is TIBER-EU?.* [Online]
Available at: https://www.ecb.europa.eu/paym/cyber-resilience/tiber-eu/html/index.en.html
[Accessed 20 July 2020].

European Commision, 2016. *General Data Protection Regulation.* [Online]
Available at: https://gdpr-info.eu
[Accessed November 2019].

European Commission, 2017. *Resilience, Deterrence and Defence: Building Strong Cybersecurity in Europe. State of the Union 2017: Cybersecurity. European Commission Fact Sheet.,* s.l.: European Commission.

European Commission, 2018. *Internal Market, Industry, Entrepreneurship and SMEs.* [Online]
Available at: http://ec.europa.eu/growth/smes/business-friendly-environment/sme-definition_en
[Accessed 30 July 2018].

European Commission, 2021. *Biometric data.* [Online]
Available at: https://ec.europa.eu/home-affairs/what-we-do/networks/european_migration_network/glossary_search/biometric-data_en
[Accessed 3 April 2021].

European Council, 2018a. *Programs. Certified Ethical Hacker Certification.* [Online]
Available at: https://www.eccouncil.org/programs/certified-ethical-hacker-ceh/
[Accessed 18 June 2018].

European Council, 2018b. *Programs. The LPT (Master) Training Program: Advanced Penetration Testing Course.* [Online]
Available at: https://www.eccouncil.org/programs/licensed-penetration-tester-lpt-master/
[Accessed 18 June 2018].

European Council, 2020. *Ethical Hacking.* [Online]
Available at: https://www.eccouncil.org/ethical-hacking/
[Accessed 14 April 2020].

European Parliament, 2016. *General Data Protection Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016.* [Online]
Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2016.119.01.0001.01.ENG&toc
[Accessed 15 June 2018].

European Parliament, 2017. *Legal Frameworks for Hacking by Law Enforcement: Identificiation, Evaluation and Comparison of Practices,* Brussels: European Parliament.

European Union, 2013. *Directive 2013/40/EU of the European Parliament and of the Council of 12 August 2013 on attacks against information systems and replacing Council Framework Decision 2005/222/JHA.* [Online]
Available at: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:218:0008:0014:EN:PDF
[Accessed 4 April 2021].

Europol, 2018. *Internet Organised Crime Threat Assessment Report (IOCTA) 2018,* s.l.: Europol.

Evan, M. & Freeman, R., 1993. A Stakeholders Theory of the Modern Corporation: Kantian Capitalism. In: *Ethical Theory and Business.* Tom L. Beachamp & Norman E Bowie (eds.) ed. Englwood Cliff's(New Jersey): Englewood Cliff's.

Evans, N. & Horsthemke, W., 2019. Active Defence Techniques. In: A. Kott & I. Linkov, eds. *Cyber Resilience of Systems and Networks.* Cham: Springer, pp. 221-224.

FBI, 2018. *The Morris Worm.* [Online]
Available at: https://www.fbi.gov/news/stories/morris-worm-30-years-since-first-major-attack-on-internet-110218
[Accessed 18 September 2020].

Ferguson, S., 2019. *Ransomware Attack Costs Norsk Hydro $40 Million - So far.* [Online]
Available at: https://www.bankinfosecurity.com/ransomware-attack-costs-norsk-hydro-40-million-so-far-a-12269
[Accessed 30 September 2019].

Ferreira, M. & Kawakami, C., 2018. Ransomware - Kidnapping personal data for ransom and the information as hostage. *ADCAIJ: Advances in Distributed Computing and Artifical Intelligence Journal,* 7(3), pp. 5-14.

Finn, P., 1995 . *The ethics of deception in research. In: Penslar RL, editor. Research ethics: Cases and materials.* Bloomington: Indiana University Press.

Freeman, R., 1984. *Strategic Management: A Stakeholder Approach..* Boston: Pittman.

Freeman, R., 1994. *The Politics of Stakeholder Theory: Some Future DIrections*. Business Ethics Quarterly*,* 4(4), pp. 409-422.

Freeman, R., 2010. *Strategic Management: A Stakeholder Approach.* Cambridge: Cambridge University Press.

Freeman, R. & Gilbert, D., 1989. *Corporate Strategy and The for Ethics.* New Jersey (Prentice Hall): Englewood Cliffs.

Freeman, R. & Gilbert, R., 1992. A Critical Agenda. Business & Society. *Business Ethics and Society,* Volume 31, pp. 9-1.

Freeman, R. et al., 2010. *Stakeholder Theory: State of the Art.* Cambridge, New York: Cambridge University Press.

Friis-Jensen, E., 2014. *The History of Bug Bounty Programs.* [Online]
Available at: https://blog.cobalt.io/the-history-of-bug-bounty-programs-50def4dcaab3
[Accessed 24 April 2020].

Fruhlinger, J., 2018. *What is ransomware? How these attacks work and how to recover from them.* [Online]
Available at: https://www.csoonline.com/article/3236183/what-is-ransomware-how-it-works-and-how-to-remove-it.html
[Accessed 4 June 2019].

F-Secure, 2018. *Incident Response Report.* [Online]
Available at: https://fsecurepressglobal.files.wordpress.com/2018/02/f-secure-incident-response-report.pdf
[Accessed 13 July 2018].

Fulton, E., Lawrence, C. & Clouse, S., 2013. W*hite hats chasing black hats: careers in it and the skills required to get there*. *J Inf Syst Educ,* 24(1), pp. 75-80.

Furnell, S. & Karweni, T., 1999. *Security implications of electronic commerce: a survey of consumers and businesses*. Electronic Networking Applications and Policy*.,* 9(5), pp. 372-382.

Gallagher, S., 2020. *New 'red team as a service' platform aims to automate hacking tests for company networks.* [Online]

Available at: https://arstechnica.com/information-technology/2020/02/the-loyal-opposition-randoris-attack-turns-red-teaming-into-cloud-service/
[Accessed 18 November 2020].

Gartner, 2015. *Gartner Says 6.4 Billion Connected "thing" Will Be in Use in 2016, Up 30 Percent From 2015.* [Online]
Available at: https://www.gartner.com/en/newsroom/press-releases/2015-11-10-gartner-says-6-billion-connected-things-will-be-in-use-in-2016-up-30-percent-from-2015
[Accessed 21 October 2020].

Gattiker, U. & Kelley, 1999. *Morality and Computers: Attitudes and Differences in Moral Judgments*. *Information Systems Research,* September, 10(3), p. 233–54.

Gelinas, L., Wertheimer, A. & Miller, F. G., 2016. *When and why is research without consent permissible? Hastings Center Report,* 46(2), p. 35–43.

General Data Protection Regulation (GDPR), 2018. *General Data Protection.* [Online]
Available at: https://gdpr-info.eu
[Accessed 21 December 2019].

Giacomello, G. & Pescaroli, G., 2019. Managing Human Factors. In: *Cyber Resilience of Systems and Networks. Risk, Systems and Decisions.* Cham: Springer, pp. 247-263.

Gilligan, C., 1982. *In A Different Voice: Psychological Theory and Women's Development.* Cambridge : Harvard University Press.

Gomez, M., 2020. *Dark Web Price Index 2020.* [Online]
Available at: https://www.privacyaffairs.com/dark-web-price-index-2020/
[Accessed 5 October 2020].

Goodpaster, K., 1991. Business Ethics and Stakeholder Analysis. *Business Ethic Quarterly,* 1(1), pp. 53-73.

Granger, S., 1994. *The Hacker Ethic.* s.l., ACM, pp. 7-9.

Greenberg, A., 2018. *The Untold Story of NotPetya, the Most Devastating Cyberattack in History..* [Online]
Available at: https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/
[Accessed Septenber 2019].

Greenburg, P., 2018. *Security Breach Notification Laws.* s.l., National Conference of State Legislatures.

GreyCampus, 2020. *Ethical Hacking.* [Online]
Available at: https://www.greycampus.com/opencampus/ethical-hacking/what-is-footprinting
[Accessed 10 June 2020].

Grimes, R., 2017. *Hacking the Hacker.* Indiana: John Wiley & Sons, Inc.

Gunarto, H., 2003. *Ethical Issues in Cyberspace and IT Society.* [Online]
Available at: http://www.apu.ac.jp/~gunarto/it1.pdf
[Accessed July 2017].

Habermas, J., 1990. Discourse Ethics: Notes on a Program of Philosophical Justification. *Moral Consciousness and Communicative Action, Christian Lenhardt Shierry Weber Nicholsen (trans.),* p. 58,66.

Hacken, n.a. *How much does penetration test cost, or price of your security.* [Online]
Available at: https://hacken.io/research/education/how-much-does-penetration-test-cost-or-price-of-your-security/
[Accessed 24 April 2020].

Hackerone, 2019. *Paypal celebrates it's first anniversary on hackerone.* [Online]
Available at: https://www.hackerone.com/blog/paypal-celebrates-its-first-anniversary-hackerone
[Accessed 24 April 2020].

Halpin, P. & Humphries, C., 2021. *Irish health service hit by 'very sophisticated' ransomware attack.* [Online]
Available at: https://www.reuters.com/technology/irish-health-service-hit-by-ransomware-attack-vaccine-rollout-unaffected-2021-05-14/
[Accessed 22 May 2021].

Hampton, N. & Baig, Z., 2015. *Ransomware: Emergence of the cyber-extortion menace.* s.l., Symantec Corporation, pp. 47-56.

Harrington, S., 1996. The Effect of Codes of Ethics and Personal Denial of Responsibility on Computer Abuse Judgments and Intentions. *MIS Quarterly,* September, 20(3), pp. 257-278.

Hart, H., 1955. Are there Any Natural Rights? *Philosophical Review,* Volume 64.

Hatfield, J., 2019. Virtuous human hacking: The ethics of social engineering in penetration-testing. *Computers & Security,* Volume 83, pp. 354-366.

Hay Newman, L., 2018. *Hacker Lexicon: What is Sinkholing?* [Online]
Available at: https://www.wired.com/story/what-is-sinkholing/
[Accessed 22 October 2020].

Henderson, T., 2017. *Short, Stealthy, Sub-Saturating DDoS Attacks Pose Greatest Security Threat to Businesses.* [Online]
Available at: https://www.missioncriticalmagazine.com/articles/91210-short-stealthy-sub-saturating-ddos-attacks-pose-greatest-security-threat-to-businesses
[Accessed 21 October 2020].

Herzog, P., 2010. *The Open Source Security Testing Methodology Manual: Contemporary Secuirty Testing and Analysis.* [Online]
Available at: https://www.isecom.org/OSSTMM.3.pdf
[Accessed 20 July 2020].

High Bit Security, 2020a. *Penetration testing - It's What We Do.* [Online]
Available at: https://highbitsecurity.com/FAQ-penetrationtesting.php
[Accessed 22 July 2020].

High Bit Security, 2020b. *Standard Penetration Test Cost Card.* [Online]
Available at: https://www.highbitsecurity.com/penetration-testing-cost.php
[Accessed 22 July 2020].

HighBit Security, n.a.. *Penetration Testing - It's what we do.* [Online]
Available at: https://highbitsecurity.com/FAQ-penetrationtesting.php
[Accessed 24 April 2020].

Himma, E. K., 2004. Targeting the Innocent: Active Defense and the Moral Immunity of Innocent Persons from Aggression. *Information, Communication & Ethics in Society,* Volume 2, pp. 31-40.

Himma, E. K. & Tavani, T. H., 2008. *The Handbook of Information and Computer Ethics.* Hoboken (New Kersey): John Wiley & Sons, Inc.

Hoffman, W. & Levite, A., 2017. *Can Active Measures Help Stabilise Cyberspace?* [Online]
Available at: https://carnegieendowment.org/2017/06/14/private-sector-cyber-defense-can-active-measures-help-stabilize-cyberspace-pub-71236
[Accessed 15 September 2020].

Holzer, C. & Lerums, J., 2016. *The Ethics of Hacking Back,* West Lafayette: CERIAS.

Homeland Security Science and Technology, 2012. *The Menlo Report: Ethical Principle Guiding Information and Communication Technology Research,* Sacramento: U.S. Department of Homeland Security Science and Technology Directorate.

Horne Cyber, 2020. *The Case for Advanced Penetration Testing: Zero-Day Vulnerabilities in Symantec ICSP.* [Online]
Available at: https://blog.hornecyber.com/executive-insights/the-case-for-advanced-penetration-testing-cve-2019-18380-zero-day-in-symantec-icsp
[Accessed 18 November 2020].

Howard, L., 2017. *2017 Cyber Risks to Intensify as Hackers Become More Cunning: Report.* [Online]
Available at: https://erisksolutions.com/2017/01/2017-cyber-risks-to-intensify-as-hackers-become-more-cunning-report/
[Accessed 18 November 2020].

IBM, P. I., 2018. *2018 Cost of a Data Breach Study: Global Overview.* [Online]
Available at:
https://databreachcalculator.mybluemix.net/assets/2018_Global_Cost_of_a_Data_Breach_Report.pdf

IDG, 2018. *IDG DDoS Report: Evolving Strategies For Handling Today's Complex & Costly Threats,* Massachusetts: A10 Networks.

Ilascu, I., 2019. *New LockerGoga Rnasomware Allegedly Used in Altran Attack.* [Online]
Available at: https://www.bleepingcomputer.com/news/security/new-lockergoga-ransomware-allegedly-used-in-altran-attack/
[Accessed 22 October 2019].

Imperva, 2019. *DDoS Report 2019: Global DDoS Threat Landscape,* San Mateo: Imperva Research Labs.

Inagaki, K., 2017. *Honda plant hit by WannaCry ransomware attack.* [Online]
Available at: https://www.ft.com/content/a0f5d047-2e20-3db9-b258-565d3be17bba
[Accessed 13 April 2020].

Infosec, 2019. *The Most Common Social Engineering Attacks [updated 2019].* [Online]
Available at: https://resources.infosecinstitute.com/common-social-engineering-attacks/#gref
[Accessed 5 February 2020].

Isaac, A., Ostroff, C. & Hope, B., 2020. *Travelex Paid Hackers Multimillion -Dollar Ransom Before Hitting New Obstacles.* [Online]
Available at: https://www.wsj.com/articles/travelex-paid-hackers-multimillion-dollar-ransom-before-hitting-new-obstacles-11586440800
[Accessed 27 July 2020].

Ivanov, A., Emm, D., Sinitsyn, F. & Pontiroli, S., 2016. *Kaspersky Security Bulletin 2016. The ransomware revolution.* [Online]
Available at: https://securelist.com/kaspersky-security-bulletin-2016-story-of-the-year/76757/
[Accessed 3 October 2019].

Jaiswal, R., 2020. *What to Do Before, During and After a DDoS Attack.* [Online]
Available at: https://thememunk.com/what-to-do-before-during-and-after-a-ddos-attack/
[Accessed 26 October 2020].

JavaPipe, 2015. *35 Types of DDoS Attacks Explained.* [Online]
Available at: https://javapipe.com/blog/ddos-types/
[Accessed 6 November 2020].

Jayaswal, V., Yurcik, W. & Doss, D., 2002. *Internet hack back: counter attacks as self-defense or vigilantism?* Raleigh, IEEE, pp. 380-386.

Johnson, M., 2013. *Cyber crime, security and digital intelligence.* New York: Routledge.

Jordan, T., 2017. A genealogy of hacking. *The International Journal of Research into New Media Technologies,* 23(5), pp. 528-544.

Kaimal, A., Unnikrishnan, A. & Namboothiri, L., 2019. Blackholing VS. Sinkholing: a Comparative Analysis. *International Journal of Innovative Technology and Exploring Engineering,* 8(7C2), pp. 15-17.

Kambourakis, G., Kolias, C. & Stavrou, A., 2017. *The Mirai Botnet and the IoT Zombie Armies.* Baltimore, MD, IEEE Military Communications Conference (MILCOM), pp. 267-272.

Kamiya, S. et al., 2018. *What is the Impact of Successful Cyberattacks on Target Firms?* s.l.: National Bureau of Economic Research.

Kashinath, T., 2021. *Command and Control Servers: Things you should know.* [Online]
Available at: https://www.secpod.com/blog/command-and-control-servers-things-you-should-know/
[Accessed 10 May 2021].

Kaspersky Lab, 2018 . *Ready… Or not: Balancing future opportunities with future risks. A global survey into attitudes and opinions on IT security.* [Online]
Available at: https://media.kaspersky.com/documents/business/brfwn/en/The-Kaspersky-Lab-Global-IT-Risk
[Accessed June 2018].

Kaspersky, 2015. *Global IT Security Risks Survey.* [Online]
Available at: http://go.kaspersky.com/rs/802-IJN-240/images/Global%20IT%20Security%20Risks%20Survey%20Ent.pdf
[Accessed 23 November 2020].

Kaspersky, 2016. *Kaspersky Security Bulletin 2016. Story of the Year: The ransomware revolution.* [Online]
Available at: https://media.kasperskycontenthub.com/wp-content/uploads/sites/43/2018/03/07182404/KSB2016_Story_of_the_Year_ENG.pdf
[Accessed 3 October 2019].

Kaspersky, 2019a. *Information security in loss figures.* [Online]
Available at: https://www.kaspersky.com/blog/security-economics-2019/28838/
[Accessed 29 January 2020].

Kaspersky, 2019b. *Story of the Year 2019: Cities under ransomware siege.* [Online]
Available at: https://securelist.com/story-of-the-year-2019-cities-under-ransomware-siege/95456/
[Accessed 7 August 2020].

Kaspersky, 2019c. *The State of Industrial Cybersecurity.* [Online]
Available at: https://ics.kaspersky.com/media/2019_Kaspersky_ARC_ICS_report.pdf
[Accessed 29 January 2020].

Kaspersky, 2020. *Kaspersky research finds ddos attacks tripled year on year in q2 2020.* [Online]
Available at: https://usa.kaspersky.com/about/press-releases/2020_kaspersky-research-finds-ddos-attacks-tripled-year-on-year-in-q2-2020
[Accessed 10 September 2020].

Kaspersky, 2021. *What is the Deep and Dark Web?* [Online]
Available at: https://www.kaspersky.com/resource-center/threats/deep-web
[Accessed 3 April 2021].

Kelly, C., Pitropakis, N., McKeown, S. & Lambrinoudakis, C., 2020. *Testing And Hardening IoT Devices Against the Mirai Botnet.* Dublin, IEE International Conference on Cyber Security and Protection of Digital Services.

Keromytis, A., 2011. *Network Bandwidth Denial of Service (DoS).* [Online]
Available at: https://www1.cs.columbia.edu/~angelos/Papers/2010/ddos.pdf
[Accessed 21 October 2020].

Kertysova, K. et al., 2018. *Cybersecurity: Ensuring awareness and resilience of the private sector across Europe in face of mounting cyber risks.* Brussels: EESC.

Kotadia, M., 2004. *Symbiot launches DDoS counter-strike tool.* [Online]
Available at: https://www.zdnet.com/article/symbiot-launches-ddos-counter-strike-tool/
[Accessed 8 October 2020].

Kouatli, I., 2016. *Managing Cloud Computing Environment: Gaining Customer Trust with Security and Ethical Management.* Amsterdam, Elsevier Science , p. 412–21.

Krebs, B., 2011. *Financial Mogul Linked to DDoS Attacks.* [Online]
Available at: https://krebsonsecurity.com/2011/06/financial-mogul-linked-to-ddos-attacks/
[Accessed 13 January 2021].

Krebs, B., 2019. *Report: No "Eternal Blue" Exploit Found in Baltimore City Ransomware.* [Online]
Available at: https://krebsonsecurity.com/2019/06/report-no-eternal-blue-exploit-found-in-baltimore-city-ransomware/
[Accessed September 2019].

Laman, A., 2019. *SANS Vulnerability Management Survey,* Unknown: SANS Institute.

Lee, R., 2015. *The Sliding Scale of Cyber Security.* [Online]
Available at: https://www.sans.org/reading-room/whitepapers/ActiveDefense/sliding-scale-cyber-security-36240
[Accessed April 2020].

Leiwo, J. & Heikkuri, S., 1998. *An Analysis of Ethics as Foundation of Information Security in Distributed Systems.* Los Alamitos, IEEE, pp. 213-22.

Lerner, Z., 2014. Microsoft the Botnet Hunter: The Role of Public-Private Partnerships in Mitigating Botnets Notes. *Harvard Journal of Law & Technology ,* 28(1), pp. 237-250.

Logan, P. & Clarkson, A., 2005. *Teaching students to hack: curriculum issues in information security. SIGCSE Bulletin,* 37(1), p. 157–161.

Lowry, P., Posey, C., Tom, R. & Bennett, R., 2014. Is Your Banker Leaking Your Personal Information? The Roles of Ethics and Individual-Level Cultural Characteristics in Predicting Organizational Computer Abuse. *Journal of Business Ethics,* 121(3), p. 385–401.

Lubis, M., Ibtisam binti Yaacob, N., binti Reh, H. & Abdulghani, M., 2011. *Study on Implementation and Impact of Google Hacking in Internet Security,* s.l.: SSRN.

Lumen, 2020. *The new cyber arms race: A changing attack landscape requires a modernised strategy.* [Online]
Available at: https://assets.centurylink.com/is/content/centurylink/lumen-new-cyber-arms-race-ddos-white-paper?Creativeid=eb33e07a-19cf-4b1a-bb3c-8c7ea86c6c36
[Accessed 3 November 2020].

MacNamee, G., 2021. *Government urged to consider paying ransom or 'doing a deal' with HSE hackers to get systems back to normal.* [Online]
Available at: https://www.thejournal.ie/cyber-attack-ransomware-5441367-May2021/
[Accessed 22 May 2021].

Malwarebytes, 2020. *Backdoor Computing Attacks: What is a backdoor?* [Online]
Available at: https://www.malwarebytes.com/backdoor/
[Accessed 4 September 2020].

Mansfield-Devine, S., 2016. DDoS goes maintsream: how headline-grabbing attacks could make this threat an organisation's biggest nightmare. *Network Security,* 2016(11), pp. 7-13.

Manuel, J. & Salvio, J., 2019. *LockerGoga: Ransomware Targeting Critical Infrastructure.* [Online]
Available at: https://www.fortinet.com/blog/threat-research/lockergoga-ransomeware-targeting-critical-infrastructure.html
[Accessed 22 October 2019].

Marble, J. et al., 2015. The Human Factor in Cybersecurity: Robust & Intelligent Defense. In: S. Jajodia, et al. eds. *Cyber Warfare, Advances in Information Security 56..* Switzerland: Springer International Publishing, pp. 173-205.

Marciano, C., 2020. *How much does Cyber/Data Breach Insurance Cost?* [Online]
Available at: https://databreachinsurancequote.com/cyber-insurance/cyber-insurance-data-breach-insurance-premiums/
[Accessed 18 January 2021].

Mari, K., 2009. *Creating sustainable work systems: developing social sustainability.* 2nd ed. New York: Routledge.

Marshall, E., 2018. *Bug bounty hunters: who are they? Why do they matter?* [Online]
Available at: https://www.mailguard.com.au/blog/bug-bounty-hunters-who-are-they
[Accessed 24 April 2020].

Mateiu, M., 2018. *The Ultmate Guide to Ransomware.* [Online]
Available at: https://www.avg.com/en/signal/what-is-ransomware
[Accessed 1 May 2019].

Mathews, L., 2017. *NotPetya Ransomware Attacks Costs Shipping Giant Maersk Over $200 Million.* [Online]
Available at: https://www.forbes.com/sites/leemathews/2017/08/16/notpetya-ransomware-attack-cost-shipping-giant-maersk-over-200-million/#770efac14f9a
[Accessed 30 September 2019].

Matwyshyn, A. M., 2009. CSR and the Corporate Cyborg: Ethical Corporate Information Security Practices. *Journal of Business Ethics,* September , Volume 88, p. 579–94.

McCallion, J., 2020. *How to protect against a DDoS attack.* [Online]
Available at: https://www.itpro.co.uk/security/ddos/28039/how-to-protect-against-a-ddos-attack
[Accessed 15 October 2020].

McGuire, M., 2018. *Into the Web of Profit: Understanding the Growth of the Cybercrime Economy,* s.l.: Bromium.

McKenzie, T., 2017. *Perspectives on Cyber Power.* [Online]
Available at: https://media.defense.gov/2017/Nov/20/2001846608/-1/-

1/0/CPP_0004_MCKENZIE_CYBER_DETERRENCE.PDF
[Accessed 11 May 2020].

McReynolds, P., 2015. How to Think About Cyber Conflicts Involving Non-State Actors. *Philosophy & Technology,* September , 28(3), p. 427–48.

Mehta, A., 2021. *'Callous' ransomware attack has caused 'catastrophic' damage to Irish health care system.* [Online]
Available at: https://news.sky.com/story/callous-ransomware-attack-has-caused-catastrophic-damage-to-irish-health-care-system-12312243
[Accessed 22 May 2021].

Merriam-Webster, 2018. *Merriam-Webster.* [Online]
Available at: https://www.merriam-webster.com/dictionary/private%20sector
[Accessed 13 July 2018].

Michael, M., 2016. *3 Surprising Things You Didn't Know About Ransomware.* [Online]
Available at: https://blog.f-secure.com/3-surprising-things-you-didnt-know-about-ransomware/
[Accessed 3 October 2019].

Miessler, D., 2016. *The Difference Between Red, Blue, and Purple Teams.* [Online]
Available at: https://danielmiessler.com/study/red-blue-purple-teams/
[Accessed 2 July 2020].

Miller, F. G., 2008. Research on medical records without informed consent. *Journal of Law, Medicine & Ethics,* 36(3), p. 560–566.

Mill, J., 1989. *On Liberty.* s.l.:Cambridge University Press.

Mimossa, M., 2014. *Facebook Carries Out Lecpetex Botnet Takedown.* [Online]
Available at: https://threatpost.com/facebook-carries-out-lecpetex-botnet-takedown/107096/
[Accessed 13 January 2021].

Mirkovic, J. et al., 2006. *Measuring Denial Of Service.* Alexandria, 2nd ACM Workshop on Quality of Protection.

Mission Critical, 2020. *The Dark Web: DDoS Attacks Sell for as Low as $10 Per Hour: Cyber economy continues to surge.* [Online]
Available at: https://www.missioncriticalmagazine.com/articles/93185-the-dark-web-ddos-attacks-sell-for-as-low-as-10-per-hour
[Accessed 5 October 2020].

Mitchell, R., Agle, B. & Wood, D., 1997. Toward a Theory of Stakeholder Identification and Salience: Defining the Prnciple of Who and What Really Counts. *Academy of Management Review,* 22(4), pp. 853-886.

Moinescu, R. et al., 2019. Aspects of human weaknesses in cyber security. *Scientific Bulletin "Mircea cel Batran" Naval Academy*, 22(1), pp. 1-9.

Moore, G., 2020. *Why You Want to Fail a Red Team Exercise.* [Online]
Available at: https://www.infosecurity-magazine.com/opinions/fail-red-team-exercise/
[Accessed 19 July 2020].

Morgan, S., 2016. *Hackerapocalyse: A Cybercrime Revelation,* s.l.: Cybersecurity Ventures.

Moulinos, K. & Pauna, A., 2013. *Good Practices for an EU ICS Testing Coordination Capability Report,* Brussels: ENISA (European Union Agency for Network and Information Security).

Mouton, F., Malan, M. M. & Venter, H. S., 2013. Social engineering from a normative ethics perspective. *Information Security for South Africa,* pp. 1-8.

Muncaster, P., 2019. *New Bedford Hit With $5.4m Ransomware Demand.* [Online]
Available at: https://www.infosecurity-magazine.com/news/new-bedford-hit-with-53m/
[Accessed 27 September 2019].

National Audit Office, 2017. *Investigation WannaCry Cyber Attack and the NHS.* [Online]
Available at: https://www.nao.org.uk/report/investigation-wannacry-cyber-attack-and-the-nhs/
[Accessed 29 April 2019].

National Cyber Security Centre, 2019. *Secure design pinciples.* [Online]
Available at: https://www.ncsc.gov.uk/collection/cyber-security-design-principles
[Accessed 4 May 2020].

Neal, P., 2019. *Protecting the Information Society: Exploring Corporate Decision Makers' Attitudes towards Active Cyber Defence as an Online Deterrence Option,* Victoria, Canada: Royal Roads University.

Net Diligence, 2019. *Net Diligence Cyber Claims Study 2019 Report.* [Online]
Available at: https://netdiligence.com/wp-content/uploads/2020/05/2019_NetD_Claims_Study_Report_1.2.pdf
[Accessed 14 January 2021].

NetScout, 2020. *NETSCOUT Threat Intelligence Report: Issue 5, 1H 2020,* Westford: NETSCOUT.

Neumann, P., 1977. *Peopleware in Systems.* Cleveland: Association for Systems management.

Newman, S., 2019. *Understanding Link Saturation Due to DDoS Attacks.* [Online]
Available at: https://www.corero.com/blog/understanding-link-saturation-due-to-ddos-attacks/
[Accessed 21 October 2020].

Nexusguard, 2020. *DDoS Threat Report 2020 Q2.* [Online]
Available at: https://blog.nexusguard.com/threat-report/ddos-threat-report-2020-q2
[Accessed 5 October 2020].

Nicholson, S., 2019. How ethical hacking can protect organisations from a greater threat. *Computer Fraud & Security,* Issue 5, pp. 15-19.

Nicho, M. & Khan, S., 2014. Identifying vulnerabilities of advanced persistent threats: an organizational perspective. *Int J Inf Secur Priv,* 8(1), pp. 1-14.

Noddings, N., 1986. *Caring, a feminine approach to ethics & moral education.* Berkeley: University of California Press.

Oakley, J., 2018. *Improving Offensive Cyber Security Assessments Using Varied and Novel Initialisation Perspectives.* Richmond, ACM.

O'Donnell, L., 2019. *N.J.'s Largest Hospital System Pays Up in Ransomware Attack.* [Online]
Available at: https://threatpost.com/ransomware-attack-new-jersey-largest-hospital-system/151148/
[Accessed 8 January 2020].

OECD, 2012. *Proactive Policy Measures by Internet Service Providers against Botnets,* Paris: OECD.

Ogu, E., Ojesanmi, O., Awodele, O. & Kuyoro, 2019. *A Botnets Circumspection: The Current Threat Landscape, and What We Know So Far.* [Online]
Available at: https://www.mdpi.com/2078-2489/10/11/337/htm
[Accessed 11 January 2021].

Opinion Matters, 2019. *Insider Data Breach survey 2019.* [Online]
Available at: https://pages.egress.com/whitepaperpage-opinionmatterinsiderthreat-05.19_downloadpage
[Accessed June 2020].

Osbourne, C., 2018. *NotPetya ransomware forced Maersk to reinstall 4000 servers, 45000 PCs.* [Online]
Available at: https://www.zdnet.com/article/maersk-forced-to-reinstall-4000-servers-45000-pcs-due-to-notpetya-attack/
[Accessed 2 October 2019].

Palmer, D., 2019. *Cybersecurity: How to get your software patching strategy right and keep the hackers at bay.* [Online]
Available at: https://www.zdnet.com/article/cybersecurity-how-to-get-your-software-patching-strategy-right-and-keep-the-hackers-at-bay/
[Accessed 13 April 2020].

Pankov, N., 2018. *Businesses and personal data: In-depth analysis of practices and risks.* [Online]
Available at: https://www.kaspersky.com/blog/data-protection-report/23824/
[Accessed 10 June 2020].

Pattinson, J., 2020. From defence to offence: The ethics of private cybersecurity. *Europeal Journal of International Security,* Volume 5, pp. 233-254.

Pearson, S., 2013. Privacy, Security and Trust in Cloud Computing. In: *Privacy and Security for Cloud Computing.* London: Springer, p. 3–42.

PECB, 2017. *Projected Cyber Attacks in 2018: A matter of when, not if?* [Online]
Available at: https://pecb.com/article/projected-cyber-attacks-in-2018---a-matter-of-when-not-if
[Accessed June 2018].

Pendse, S. G., 2011. Ethical Hazards: A Motive, Means, and Opportunity Approach for Curbing Corporate Unethical Behavior. *Journal of Business Ethics,* 107(3), p. 265–279.

Perlroth, N., 2021. *Irish hospitals are latest to be hit by Ranswaomre Attacks.* [Online]
Available at: https://www.nytimes.com/2021/05/20/technology/ransomware-attack-ireland-

hospitals.html
[Accessed 22 2021 May].

Perlroth, N. & Isaac, M., 2018. *Insider Uber's $100,000 payment to a Hacker, and the Fallout.* [Online]
Available at: https://www.nytimes.com/2018/01/12/technology/uber-hacker-payment-100000.html
[Accessed 15 June 2018].

Pfleeger, C., Pfleeger, S. & Margulies, J., 2015. *Security in Computing.* 5th ed. Massachusetts: Pearson Education.

Phillips, R., 2003a. Stakeholder Legitimacy. *Business Ethics Quarterly,* 13(1).

Phillips, R., 2003b. *Stakeholder Theory and Organizational Ethics.* 1st ed. San Francisco: Berrett-Koehler Publishers Inc.

Phillips, R. A., 1997. Stakeholder Theory and a Principle of Fairness. *Business Ethics Quarterly,* 7(1), pp. 51-66.

Phillips, R., Freeman, E. & Wicks, A., 2003. What Stakeholder Theory Is Not. *Business Ethics Quarterly,* 13(4), pp. 479-502.

Pienta, D., Thatcher, J., Sus, H. & George, J., 2018. *Information systems betrayal: When cybersecurity systems shift from agents of protection to agents of harm.* [Online]
Available at: https://core.ac.uk/download/pdf/301376192.pdf
[Accessed 13 November 2020].

Pieters, W., 2011. Security and Privacy in the Clouds: A Bird's Eye View. In: *In Computers, Privacy and Data Protection: An Element of Choice.* s.l.:Springer, p. 445–457.

Plohmann, D., Gerhards-Padilla, E. & Leder, F., 2011. *Botnets: Detection, Measurement, Disinfection & Defence,* Brussels: ENISA.

Ponemon Institute, 2018. *Ninth Annual Cost of Cybercrime Study: Unlocking the value of improved sybersecurity protection,* s.l.: Accenture.

Ponemon Institute & IBM, 2018. *2018 Cost of a Data Breach Study: Global Overview.* [Online]
Available at: www.ibm.com/downloads/ cas/861MNWN2
[Accessed June 2020].

Posey, C., Bennett, B., Roberts, T. & Lowry, P., 2011. *When Computer Monitoring Backfires: Invasion of Privacy and Organizational Injustice as Precursors to Computer Abuse.* [Online]
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1958530.
[Accessed July 2017].

PwC(UK), 2018. *Global State of Information Security Survey 2018. Protecting digital society from cyber shocks: how prepared are UK organisations?.* [Online]
Available at: Available from: https://www.pwc.co.uk/issues/cyber-security-data-privacy/insights/global-state
[Accessed 18 June 2018].

Radware, 2017. *History of DDoS Attacks.* [Online]
Available at: https://security.radware.com/ddos-knowledge-center/ddos-chronicles/ddos-attacks-history/
[Accessed 23 October 2020].

Rawls, J., 1964. *Legal Obligation and the Duty of Fair Play in S.Hook (ed.).* New York: New York University Press.

Rawls, J., 1971/1999. *The Independence of Moral Theory in Samuel Freeman (ed.).* Cambridge(MA): Harvard University Press.

Rawls, J., 1971. *A Theory of Justice.* Cambridge(MA): The Belknap Press of Harvard University Press.

Redscan, 2020. *Whether acting out of malice or negligence, insider threats pose a significant risk to all organisations.* [Online]
Available at: https://www.redscan.com/news/a-guide-to-insider-threats-in-cyber-security/
[Accessed 29 June 2020].

Resnik, D. & Finn, P., 2017. *Ethics and Phishing Experiments,* New York: Springer.

Revuelto, V., Meintanis, S. & Socha, K., 2017. *CERT-EU Security Whitepaper 17-003: DDoS Overview and Response Guide.* [Online]
Available at: https://cert.europa.eu/static/WhitePapers/CERT-EU_Security_Whitepaper_DDoS_17-003.pdf
[Accessed 28 October 2020].

Rifaut, A., Feltus, C., Turki, S. & Khadraoui, D., 2015. *Analysis of the Impact of Ethical Issues on the Management of the Access Rights.* s.l., ACM, pp. 12-19.

Rivero Lopez, M., 2019. *LockerGoga Ransomware Family Used in Tagreted Attacks.* [Online]
Available at: https://www.mcafee.com/blogs/other-blogs/mcafee-labs/lockergoga-ransomware-family-used-in-targeted-attacks/
[Accessed 3 January 2020].

Robertson, C., Lamin, A. & Livanis, G., 2010. Stakeholder Perceptions of Offshoring and Outsourcing: The Role of Embedded Issues. *Journal of Business Ethics,* August , 95(2), pp. 167-189.

Russell, G., 2017. Resisting the persistent threat of cyber-attacks. *Computer Fraud & Security,* 2017(12), pp. 7-11.

Saha, S., 2020. Advances in Intelligent Systems and Computing. *Advances in Intelligent Systems and Computing,* Volume 1065, p. 203.

Salahdine, F. & Kaabouch, N., 2019. Social Engineering Attacks: A Survey. *Future Internet,* 11(4), p. 89.

Salman, A., Saad, S. & Ali, M. N. S., 2013. Dealing with Ethical Issues among Internet Users: Do We Need Legal Enforcement? *Asian Social Science,* 9(8), p. 3.

Sandia National Laboratory, 2011. *The Information Design Assurance Red Team (IDART).* [Online]
Available at: http://idart.sandia.gov
[Accessed 7 September 2020].

Sarhan, A., Farhan, S. & Al-Harby, F., 2018. Understanding and discovering sql injection vulnerabilities. *Advances in Intelligent Systems and Computing,* Issue 593, pp. 45-51.

Schoeman, F., 1984. *Philosophical Dimensions of Privacy: An Anthology.* 3 ed. Cambridge: Cambridge University Press.

Secarma, 2020. *Red Team Assessment.* [Online]
Available at: https://www.secarma.com/services/red-teaming-services/red-team-assessment.html
[Accessed 29 June 2020].

Shackelford, S., Charoen, D., Waite, T. & Zhang, N., 2019. Rethinking Active Defence: A Comparative Analysis of Proactive Cybersecurity Policymaking. *University of Pennsylvania Journal of International Law,* 41(2), pp. 377-427.

Shakib, J. & Layton, D., 2014. *Interaction between Ethics and Technology.* s.l., IEEE.

Shearwater, 2017. *Vulnerability Management 101: 5 Best Practices for Success.* [Online]
Available at: https://www.shearwater.com.au/vulnerability-management-ebook/
[Accessed 1 May 2020].

Simshaw, D. & Wu, S., 2015. *Ethics and Cybersecurity: Obligations to Protect Client Data.* San Francisco, CA, s.n.

Singapore CERT, 2020. *Ransom Distributed Denial-of-Service Attacks.* [Online]
Available at: https://www.csa.gov.sg/singcert/publications/ransom-distributed-denial-of-service-attacks
[Accessed 6 November 2020].

Sjouwerman, S., 2018. *MSPs: Ransomware Downtime Costs SMBs 10x the Ransom.* [Online]
Available at: https://blog.knowbe4.com/msps-ransomware-downtime-costs-smbs-10x-the-ransom
[Accessed 30 September 2019].

Sloane Risk Group, 2020. *Red Teaming.* [Online]
Available at: https://sloaneriskgroup.co.uk/red-teaming/
[Accessed 18 November 2020].

Smart, W., 2018. *Lessons learned review of the WannaCry Ransomware Cyber Attack.* [Online]
Available at: https://www.england.nhs.uk/wp-content/uploads/2018/02/lessons-learned-review-wannacry-ransomware-cyber-attack-cio-review.pdf
[Accessed June 2018].

Sophos, 2021. *The State of Ransomware 2021.* [Online]
Available at: https://secure2.sophos.com/en-us/content/state-of-ransomware.aspx
[Accessed May May 2021].

Spadafora, A., 2020. *Travelex forked out multi-million ransom to restore its systems.* [Online]
Available at: https://www.techradar.com/nz/news/travelex-forked-out-multi-million-ransom-to-

restore-its-systems
[Accessed 27 July 2020].

Spafford, E., 1989. *The Internet Worm Incident.* Indiana: Springer -Verlag.

Stahl, B., Doherty, N., Shaw, M. & Janicke, H., 2014. Critical theory as an approach to the ethics of information security. *Sci Eng Ethics ,* 20(3), p. 675–99.

Starik, M., 1995. Should Trees Have Managerial Standing? Toward Stakeholder Status for Non-Human Nature. *Journal for Business Ethics,* Volume 14, pp. 207-217.

Statista, 2018. *Internet of Things (IoT) active device connections installed base worldwide from 2015 to 2025.* [Online]
Available at: https://www.statista.com/statistics/1101442/iot-number-of-connected-devices-worldwide/
[Accessed 24 September 2020].

Stevens, S., 2020. A Framework For Ethical Cyber-Defence for Companies. In: M. Christen, B. Gordijn & M. Loi, eds. *The Ethics of Cybersecurity.* Cham: Springer, pp. 317-330.

Stiawan, D. et al., 2017. Cyber-Attack Penetration Test and Vulnerability Analysis. *International Journal of Online and Biomedical Engineering,* 13(1), pp. 125-132.

Stuart Mill, J., 1989. *On Liberty.* Cambridge: Cambridge University Press.

Sucuri, 2019. *What is a DDoS Attack?* [Online]
Available at: https://sucuri.net/guides/what-is-a-ddos-attack/
[Accessed 18 September 2020].

Sussman, B., 2019. *Special Security Advisory: 'Ryuk Ransomware Targeting Organisations Globally'.* [Online]
Available at: https://www.google.com/amp/s/www.secureworldexpo.com/industry-news/how-ryuk-ransomware-works%3fhs_amp=true
[Accessed 3 October 2019].

Symanovich, S., n.d. *5 Reasons why general software updates and patches are important.* [Online]
Available at: https://us.norton.com/internetsecurity-how-to-the-importance-of-general-software-updates-and-patches.html
[Accessed 12 April 2020].

Symantec, 2018. *2018 Internet Security Threat Report,* s.l.: Symantec.

Symantec, 2019. *Targeted Ransomware: Proliferating Menace Threatens Organizations.* [Online]
Available at: https://www.symantec.com/blogs/threat-intelligence/targeted-ransomware-threat
[Accessed 22 October 2019].

Taddeo, M., 2013. Cyber Security and Individual Rights, Striking the Right Balance. *Philosophy and Technology,* 26(4), pp. 353-356.

Taddeo, M., 2015. The Struggle Between Liberties and Authorities in the Information Age. *Science and Engineering Ethics,* October, 21(5), p. 1125–38.

Tavani, H., 2013. *Ethics & Technology. Controversies, Questions, and Strategies for Ethical Computing.* 4th ed. s.l.:Wiley.

Technopedia, n.d.. *What is a black hat hacker? - defintition from Technopedia.* [Online]
Available at: https://technopedia.com/definition/26342/black-hat-hacker
[Accessed 7 July 2019].

Tennessee, U. o., n.d. *Office of Information Technology.* [Online]
Available at: https://help.utk.edu/kb/index.php?func=show&e=1960
[Accessed 22 October 2019].

The Australian Cyber Security Centre (ACSC), et al., 2018. *Joint Report on publicly available hacking tools.* [Online]
Available at: https://www.ncsc.gov.uk/report/joint-report-on-publicly-available-hacking-tools
[Accessed 4 November 2019].

Thomas, A., 2017. *Germany's Deutsche Bahn Rail Operator Targeted in Global Cyberattack.* [Online]
Available at: https://www.wsj.com/articles/germanys-deutsche-bahn-rail-operator-targeted-in-global-cyberattack-1494658493
[Accessed 13 April 2020].

Thomas, G., Low, G. & & Burmeister, O., 2018. "Who Was That Masked Man?" : System Penetrations - Friend or Foe. In: H. Prunckun, ed. *Cyber Weaponry, Advanced Sciences and Technologies for Security Applications.* Cham: Springer, p. 122.

Thomas, K., 2015. *Ransomware: Should you pay the criminals.* [Online]
Available at: https://www.welivesecurity.com/2015/04/23/ransomware-pay-cybercriminals/
[Accessed 3 October 2019].

Thompson, C., 2019. *Penetration Testing Versus Red Teaming: Clearing the Confusion.* [Online]
Available at: https://securityintelligence.com/posts/penetration-testing-versus-red-teaming-clearing-the-confusion/
[Accessed 26 May 2020].

Tidy, J., 2019. *How a ransomware attack cost one firm £45m.* [Online]
Available at: https://www.bbc.com/news/business-48661152
[Accessed June 2019].

Tripwire, 2015. *Tripwire. 11 Essential Bug Bounty Programs of 2015.* [Online]
Available at: https://www.tripwire.com/state-of-security/vulnerability-management/11-essential-bug-bounty-programs-of-2015/
[Accessed 18 June 2018].

TSA, 2020. *Global DDoS Threat Report 2019.* [Online]
Available at: https://www.trustwave.com/en-us/resources/library/documents/global-ddos-threat-report/
[Accessed 6 November 2020].

Upadhyaya, R., 2016. *Cyber Ethics and Cyber Crime: a deep dwelved study into legality, ransomware, underground web and bitcoin wallet.* Noida, India, IEEE, pp. 143-148.

US-CERT, 2012. *ICS-CERT Monthly Monitor.* [Online]
Available at: https://us-cert.cisa.gov/sites/default/files/Monitors/ICS-CERT_Monitor_Aug2012.pdf
[Accessed 17 September 2020].

van Eeten, M., Bauer, J. A. H. & Tabatabaiei, S., 2010. *The Role of Internet Service Providers in Botnet Mitigation an Empirical Analysis Based on Spam Data.* Twente: OECD.
Varonis, 2017. *Ransomware Defense Survey 2017: The Enterprise Strikes Back,* s.l.: SMG Information Security Media Group.

Venkatraman, S. & Delpachitra, I., 2008. Biometrics in banking security: a case study. *Information Management & Computer Security,* 16(4), pp. 415-430.

Vigliarolo, B., 2017. *10 major organizations affected by the WannaCry ransomware attack.* [Online]
Available at: https://www.techrepublic.com/pictures/gallery-10-major-organizations-affected-by-the-wannacry-ransomware-attack/9/
[Accessed 13 April 2020].

Vishwakarma, R. & Jain, A., 2019. A survey of DDoS attacking techniques and defence mechanisms in the IoT network. *Telecommunication Systems,* Volume 73, pp. 3-25.

Walker, J., 2018. *Ransomware remains biggest malware threat in 2018, says Europol.* [Online]
Available at: https://portswigger.net/daily-swig/ransomware-remains-biggest-malware-threat-in-2018-says-europol
[Accessed 1 May 2019].

Walters, G. J., 2001. Privacy and Security: An Ethical Analysis. *SIGCAS Comput. Soc.* June, 31(2), pp. 8-23.

Watkins, L., Silberberg, K., Morales, J. & Robinson, W., 2015. *Using Inherent Command and Control Vulnerabilities to halt DDoS Attacks.* s.l., IEEE, pp. 3-10.

Wenger, F. & et al, 2017. *Canvas White Paper 3 – Attitudes and Opinions Regarding Cybersecurity.* [Online]
Available at: https://ssrn.com/abstract=3091920
[Accessed June 2018].

Whittaker, Z., 2020. *Red teams OK to push ethical limits but not on themselves, study says.* [Online]
Available at: https://techcrunch.com/2020/02/02/red-team-ethical-limits/
[Accessed 19 July 2020].

Wilhelm, T., 2013. *Professional Penetration Testing.* 2nd ed. s.l.:Elsevier Inc.

Williams, R., 2015. *Internet Security Made Easy: Take Control of Your Online World.* London: Flame Tree Publishing.

World Economic Forum, 2020. *The Global Risks Report 2020,* s.l.: World Economic Forum.

Wright, S., 2018. *https://blog.nettitude.com/a-red-teaming-approach-to-pci-dss.* [Online]
Available at: A red teaming approach to PCI-DSS
[Accessed 18 November 2020].

Yaghmaei, E. et al., 2017. *Canvas White Paper 1 – Cybersecurity and Ethics.* [Online]
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3091909
[Accessed June 2018].

Yimu, J. & Shangdong, L., 2019. *Threats for Botnets.* [Online]
Available at: https://www.intechopen.com/books/computer-security-threats/threats-from-botnets
[Accessed 31 December 2020].

Zetter, K., 2011. *With Court Order, FBI Hijacks 'Coreflood' Botnet, Sends Kill Signal.* [Online]
Available at: http://www.wired.com/2011/04/coreflood
[Accessed 24 December 2020].

Zetter, K., 2015. *Hacker Lexicon: Botnets, the Zombie Computer Armies That Earn Hackers Millions.* [Online]
Available at: https://www.wired.com/2015/12/hacker-lexicon-botnets-the-zombie-computer-armies-that-earn-hackers-millions/
[Accessed 23 October 2020].